(continued from part 38)

Input methods
There are many methods of locating a point the screen and most involve moving a cross, designated by its X, Y co-ordinates, around the screen.

One method of doing this that we have seen before is the **light pen**. This is a photo-sensitive device, the size of a fountain pen, which is connected to a computer. When held close to the screen, the photo-sensor in the light pen's nib or ball detects the position on the screen and sends this information to the computer. A cross then appears at that point.

Other input devices actually move a cross around the screen, for example, the **mouse** and **tracker ball**. Mice are small, hand-held devices about the size of a bar of soap. Set into the bottom of each mouse is a rolling ball. Moving the mouse around the surface of a desk moves the ball. Sensors detect this movement and move the cross around the screen accordingly. The mouse gets its name from its motion and its tail – the lead that connects it with the computer.

Tracker balls also work on this principle but, in this case, the ball is set into the surface of the operators workstation. **Joysticks**, familiar to video game users, can also be employed, but these do not have the accuracy of other methods.

Digitizing tablets or **graphics tablets** enable designers to incorporate an existing diagram into a computer based drawing. The digitizing tablet looks rather like a normal drawing board that is wired up to a computer. A **stylus** or **cross-hair cursor**, which is sometimes known as a **puck**, is used to identify the key features and points on the drawing. Drawings are placed on the surface of the digitizing tablet and the lines and points are followed with the stylus or cursor – reproducing the original drawing on the screen. Even with no drawing present, a digitizing tablet can still be used to indicate a point on the screen and thus create a drawing.

Commands
A designer using a CAD system has to do more than simply locate and move points on a CRT screen. Lines of different types have to be created, accurate radii and curves drawn, angles set and lettering added. All of these operations are controlled by the CAD system's commands or instructions.

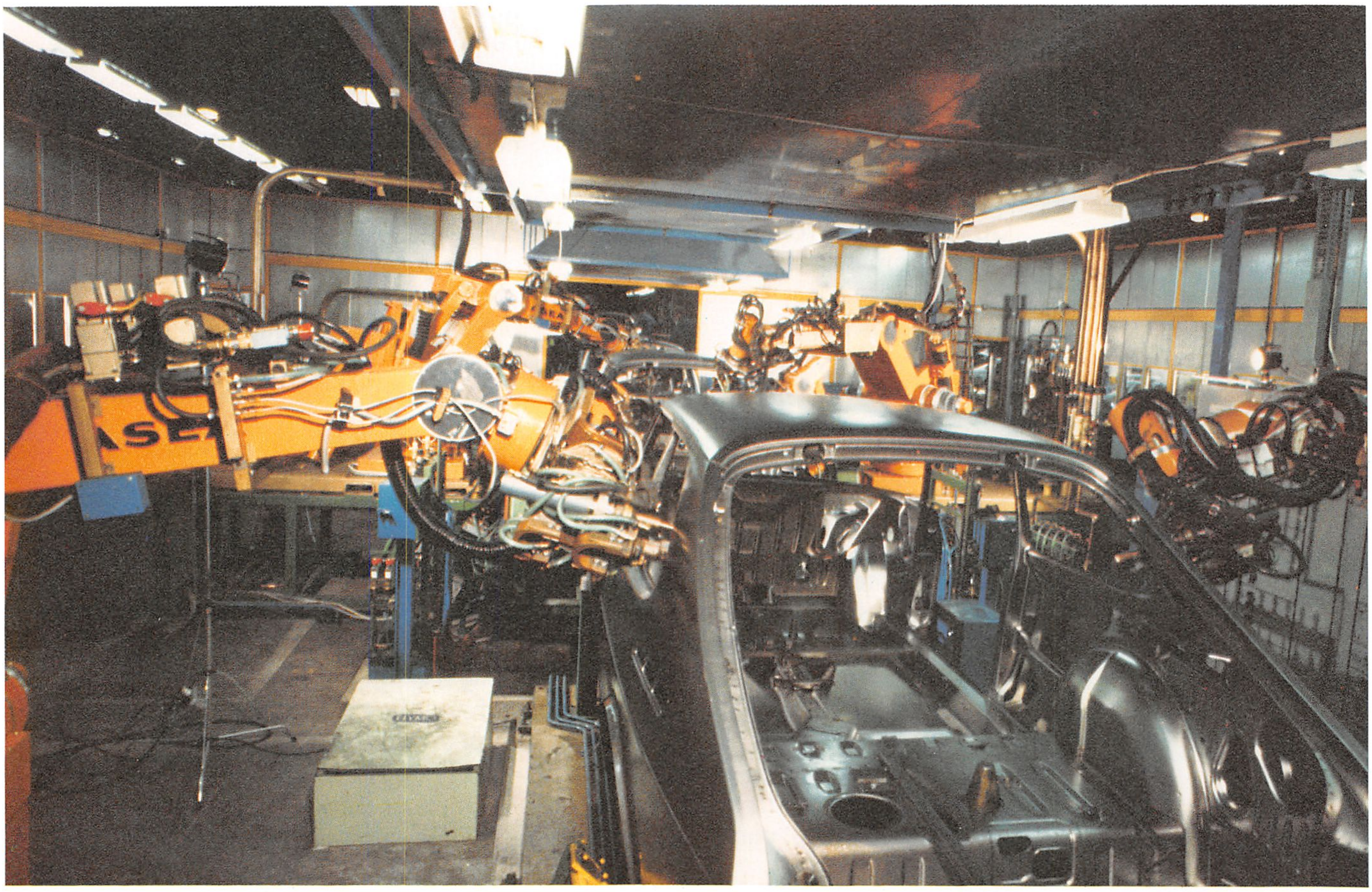
Commands can be issued to the system in different ways, according to the hardware supplied. Sometimes an ordinary keyboard is used, however, as you might imagine, this is a rather slow process. Most systems, therefore, either use a special keyboard with one button for each function, or utilise some method of function choice. For example, a digitizing tablet might have a list of functions arranged down one side – functions are selected by touching the appropriate button with the stylus.

Similarly, a CRT display might have a menu of commands displayed along the top or bottom of the screen, to be chosen using the light pen. A conventional alphanumeric keyboard is necessary to input the lettering needed on the drawing.

The commands available to the CAD system user generally relate to the kind of line to be drawn and the shape it is to take. The line might be thick, thin, dotted, chained or pecked; the shape might be a straight line linking two points; a circle linking three points; an arc tangential to two lines to form a smooth corner; a spline providing a smooth line through a number of points; a rectangle with the diagonal given by two points; and so on.

Commands can also reference frequently used shapes and symbols held in libraries. An electronics circuit design program, for example, will contain the symbols for all components, while a mechanical engineering CAD package would include a library of fastner shapes – nuts, bolts and so on. Various founts of lettering would also be available, in different type sizes and styles, so that lettering, dimensions and other annotation can be added.

Drawings can be amended by commands offering selective erasure. The screen can also act like a 'window' on the drawing enabling the operator to 'pan and tilt' around the picture and 'zoom' in and out on specific details. Drawings can then be filed and output to paper via a computer operated pen plotter.



The Research House/ASEA

Above: car welding by robot. Each robot is programmed to perform a specific sequence of actions in a certain time as the car body moves along the line.

Benefits of CAD

One of the aims of CAD systems is to relieve the designer and draughtsperson from the large amount of work that goes into handmade drawings. However, CAD systems are tools to be used by people – they cannot replace the individual. The creativity involved in drawing and design is still very important.

The main benefit of CAD over traditional draughting methods is that modifications can be made to drawings in much less time as erasing and redrawing is clean and immediate. The actual creation of drawings is also quicker because special tools and templates do not have to be used.

The computer can also perform the arithmetic needed for dimensioning components. In some systems, the calculated dimensions are added to the drawing automatically, in others the computer checks the figures provided by the draughtsperson.

Computerised filing systems can also help in the filing and retrieval of drawings – another factor that adds to CAD's ability to speed up the design process.

The reduction of draughting time is a very important factor in today's industry. The drawing stage of a product has traditionally been a bottleneck, with two thirds of the designers' and draughtspersons time being taken up by actual drawing – only one third of their time being occupied by all their other functions, including design.

Of course, new skills have to be learnt by draughting personnel and the high equipment cost involved can sometimes lead to the introduction of shift work so that expensive hardware is fully utilised. Neither of these two social effects will be willingly accepted by everyone in one company. There is the understandable fear of the unknown, especially when it is related to the possible threat to one's livelihood.

Bearing this in mind, most companies running CAD systems introduce them gradually, operating them in parallel with the traditional draughting department.

Linking design to manufacture

The real and immediate benefits of CAD systems appear when the draughting system is linked to other parts of the design and manufacturing process. The drawings of a product represent a description that can be interpreted and used by a designer for visualising and analysing that product.

From the design room to production planning, sales and buying departments, drawings are used in each stage of the manufacturing process. It follows therefore, that CAD has another great value – the computerised data that goes to make up the drawing can be used by other departments, as part of the prototyping and testing process, or to drive machine tools, or to plan a marketing campaign.

The automatic testing of a design by **simulation** is one such process. We'll take as an example aircraft wing design. The aerodynamic design of the wing would probably be based on earlier proven work. Once the basic shape is established the computer can be used to optimise the weight and strength ofthe components. This achieved, the design can be run through computer programs to calculate stress patterns in the components. The wing will also be heated by its movement through the air, so other programs can be written to calculate the distribution of heat and the effect of temperature on strength, for example.

Once the design is finalised, CAD systems can also be used to generate the data needed by the machines that machine the various components.

Products with less of a technological emphasis are also benefiting from integrated CAD systems. When a shoe is designed, for example, it has to be manufactured in a range of sizes. Using traditional methods, the dimensions of the leather parts that go to make up each size shoe had to be worked out by hand. Now a computer can automatically work out the different sizes necessary, from one drawing. The various manufacturing processes involved can also be carried out automatically.

Complex three-dimentional colour graphics packages now enable shoes, and other products, to be visualised before they are made.

CAM and CNC

The same basic industrial manufacturing, machining and shaping operations have been in use for over 100 years. These include: cutting operations, such as turning, drilling, shaping, milling; forming operations, for example embossing, rolling, extruding and bending; and shearing operations, like punching and guillotining. More recently, flame cutters, laser beams, controlled corrosion, spark corrosion, and ultrasonic and plasma arc machining have been introduced. And that's only in metal working!

The contribution that computers have made to these processes has been in terms of the precise control of the tool and workpiece movements in a programmed sequence of actions that is now possible, and the ability of the computer to file data about the machining of different parts.

Computer numerically controlled (CNC) machine tools are the devices that carry out these operations. These form part of the arsenal of **computer aided manufacturing** (CAM) devices available to industry today. Some CAM machines can be driven directly from CAD created data, thus reducing time and costs.

Component manufacture

Data about the shape of a component to be manufactured is programmed in a language specially developed to describe geometric information. The program includes a description of the tools to be used and their operating conditions (such as the speed of rotation of the tool) and indicates the pattern of movement that the machining tool has to follow.

The programs are translated by computer into the sequence of tool and workpiece movements necessary to make the component. These programs can be held on tape but machine tools are increasingly directly linked to computers that hold the prepared programs in their memories. The link between the computer and the machine tool is often two way, with the machine tool reporting its condition and the computer responding with instructions.

Machine tools contain their own computers to interpret the programs,

provide the feedback control, control the changing of tools, supervise safety systems and monitor temperatures, utilization, currents, voltage, forces, vibrations and tool wear. The machine tools' automatic sequencing and control enables them to run unattended.

Under utilization of machinery is one of the major shop floor inefficiencies. Improvements in cutting speeds made possible by new tool materials have worsened the operating time to idling time ratio. New tools have reduced machining times, but the loading and unloading of workpieces takes substantially the same time that it has always done.

Machine tools can now spend a great deal of time simply waiting: waiting for new raw materials; for someone to load or unload new parts; for replacement tools; or waiting to be serviced.

Much of the development effort that has been put into recent CAM systems has been aimed at improving machine utilization. This has been achieved by reducing machine specialization, so that each machine now performs many different tasks.

The first step in reducing specialization is to simplify the transfer of new programs to a machine tool, and the second step is to increase the range of tasks that one machine can perform. The resultant general purpose machine tool forms a part (or the whole) of what is known as a **flexible manufacturing system**. The principle features of such a machine are its ability to handle a range of tools, and to change them under the control of a program.

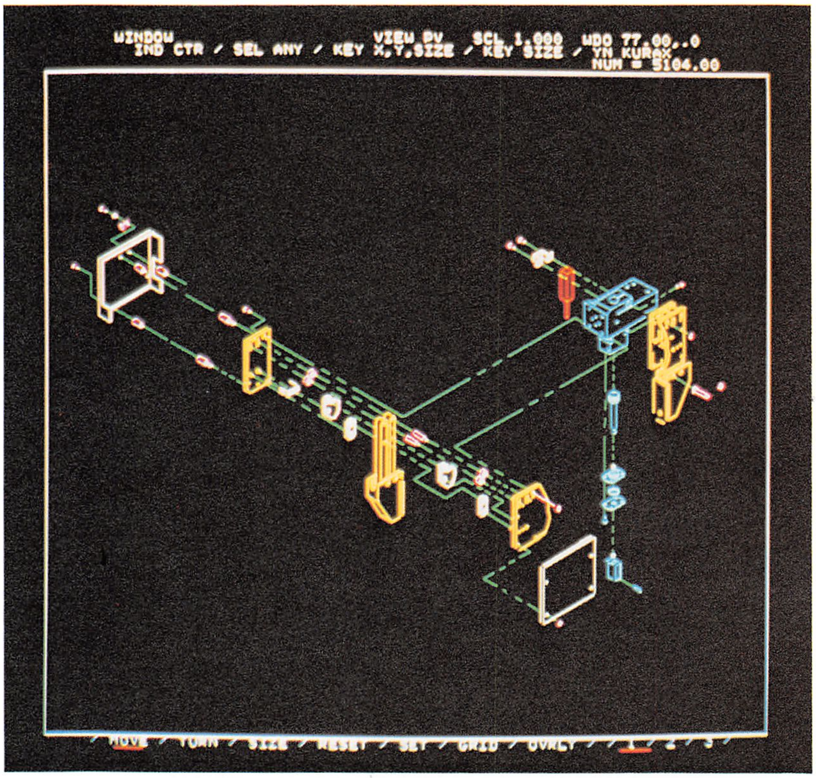
Quality control

In spite of all the precautions that can be taken during the manufacturing process, some components will inevitably prove faulty. **Quality control** ensures that the manufactured parts are monitored for faults; and if faults are found, the cause identified and corrected.

The manual measurement and record keeping involved in this process is carried out on the shop floor, with the calculations being performed later in an office. The delays inevitably incurred in this process mean that the detection of faults is slow and during this time faulty parts continue to be made.

Computers can obviously be of help in performing the statistical calculations, and some measuring instruments such as electronic micrometers can be connected directly to a computer system. The computer then automatically records the measurements that an operator takes,

Below: a multicolour CAD VDU showing an exploded view of a newly designed component. Its constituent parts and method of construction can be easily seen. (Photo: IBM).

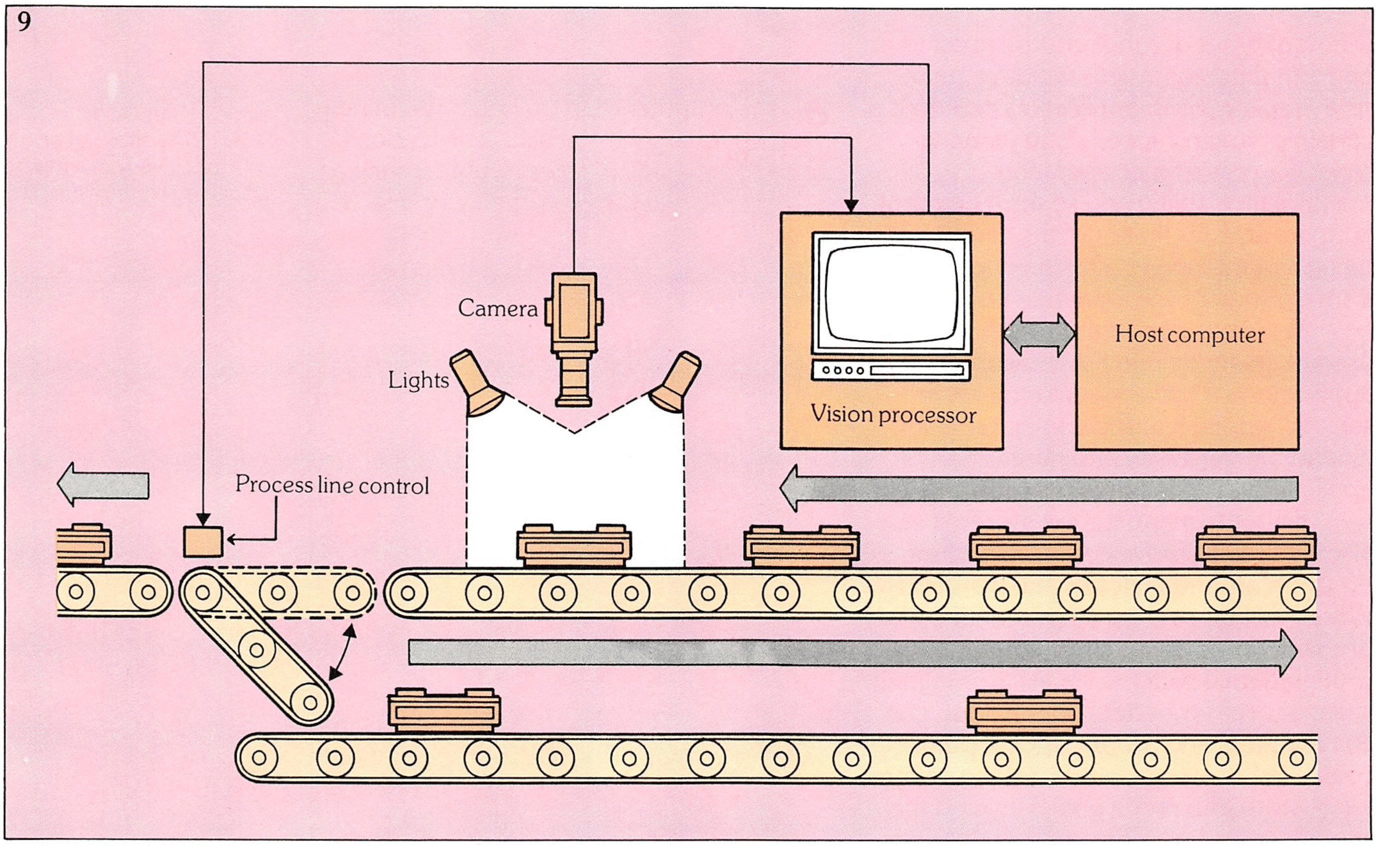


calculates the statistics, displays the results and prints a report that can accompany the goods to the customer, and act as a management record.

Developments in **vision systems**, and other kinds of sensors, now mean that it is possible to inspect and measure components without removing them from the machine tool or the conveyer. The speed of automatic measurement makes it possible to measure every single part made. This means that there is no need to rely on statistical estimates and the quality of each component that passes inspection is thus assured. The operation of a vision system is shown in *figure 9*.

Vision systems can be considered to be general purpose, in as much as they can monitor many possible dimensions and shapes – and even surface finishes. These

9



9. Automated vision systems enable quality control checking at all stages of the production process.

systems are computer based – necessarily so, as their operation depends on complex calculations. If the computer stores the characteristics of a number of components, then the system can check them as they are made. A vision system can therefore add automatic inspection and rejection to a flexible manufacturing system.

Some products use components that are known to have ill defined characteristics. In these cases, designers build means of adjustment into the product. Some clocks, for instance, use electronic motors which regulate the hands through a set of gears. Electronic machinery on the production line can monitor the oscillator to check whether the clock is fast or slow. Automatically driven screwdrivers then adjust the oscillator correcting any deviations in the time keeping. More sophisticated methods of inspection and adjustment of electronic and mechanical products can also be provided.

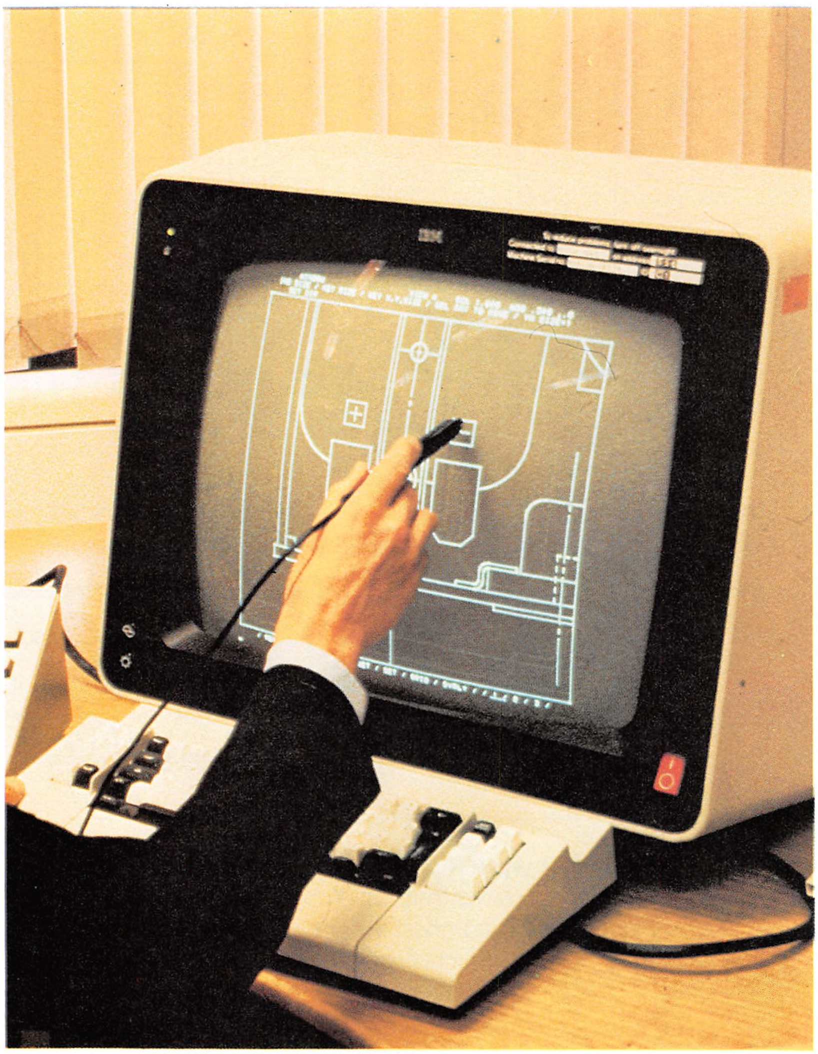
The performance of machine tools can be monitored by sensors. Records kept on computer allow the selective examination of collected data and when things go wrong, can prove invaluable in diagnosing faults. The rate of tool wear can be derived from such information – indicating when sharpening or replacement are necessary. Trouble spots on the shop floor will be highlighted, indicating when preventative servicing and maintenance are needed.

Assembly

The assembly of components into finished products is, of course, an important part of manufacturing industry. Assembly processes can be carried out by people, or by machines. Only large-scale manufacturing industries have traditionally been able to benefit from specially built assembly machines, because of the high costs involved.

Programmable robot arms offer an alternative to conventional machinery. A computer operating a robot arm provides the possibility of describing an assembly task in a textural form – this can be interpreted by the computer and converted into suitable arm movements.

Bear in mind though, that the robot is not a completely general purpose device.



Above: designing a component on the screen with the aid of a light pen.
(Photo: IBM).

Different types of robot are constructed to. perform different tasks – welding, paint spraying, etc. Each application demands a certain operating speed, load carrying capacity, precision of operation, manipulative ability and so on. The robot grippers and jigs that hold the components being assembled are often limited to quite a small range of components, for instance.

However, robots are now widely used in assembly and manufacturing applications. Further developments in sensors, grippers and vision systems will extend the versatility and range of applications of the industrial robot.

Robot transportation systems can also be used to move components and products around factories and warehouses. Where distances are greater than the reach of a robot arm, robot carts, communicating with guidance systems via electromagnetic links to wires embedded in the floor, can convey parts and products. Computer controlled conveyors and stacking systems fetch, store and move goods around the shop floor and warehouse. If shelves become full, or blockages occur, then the computer can automatically re-route the goods when sensors detect problems.

Group technology

Many of the changes that have taken place in manufacturing industry have been in the way that work is organised on the shop floor. One such example is **group technology**, where components or products with similar characteristics are handled in the same factory area by the same team of people.

The aim of this is to produce specialist units that can operate efficiently. A unit working with few machines or materials, or with few component shapes, should quickly gain experience and learn how to do a better job or overcome new problems.

The classification of each group's jobs can be based on a number of different criteria. All the cylindrical shaped components of all the factory's products may be manufactured together, for example. There are several ways in which computers can make such a scheme practicable.

Each group on the shop floor will be making components for a variety of products, so a computerised system is essential to keep a record of the location of each products' components, and what progress they are making through the various groups.

If production has to be stepped up, then a number of unconnected groups will be affected. The computer identifies them and updates their individual work plans.

Computer classification of components also allows designers to search for components already in production that are similar to those specified in a new design. If a similar component exists, then the design can be amended to use that standard item. Component classification also helps product planners to ensure a smooth flow of parts through the factory.

Integration

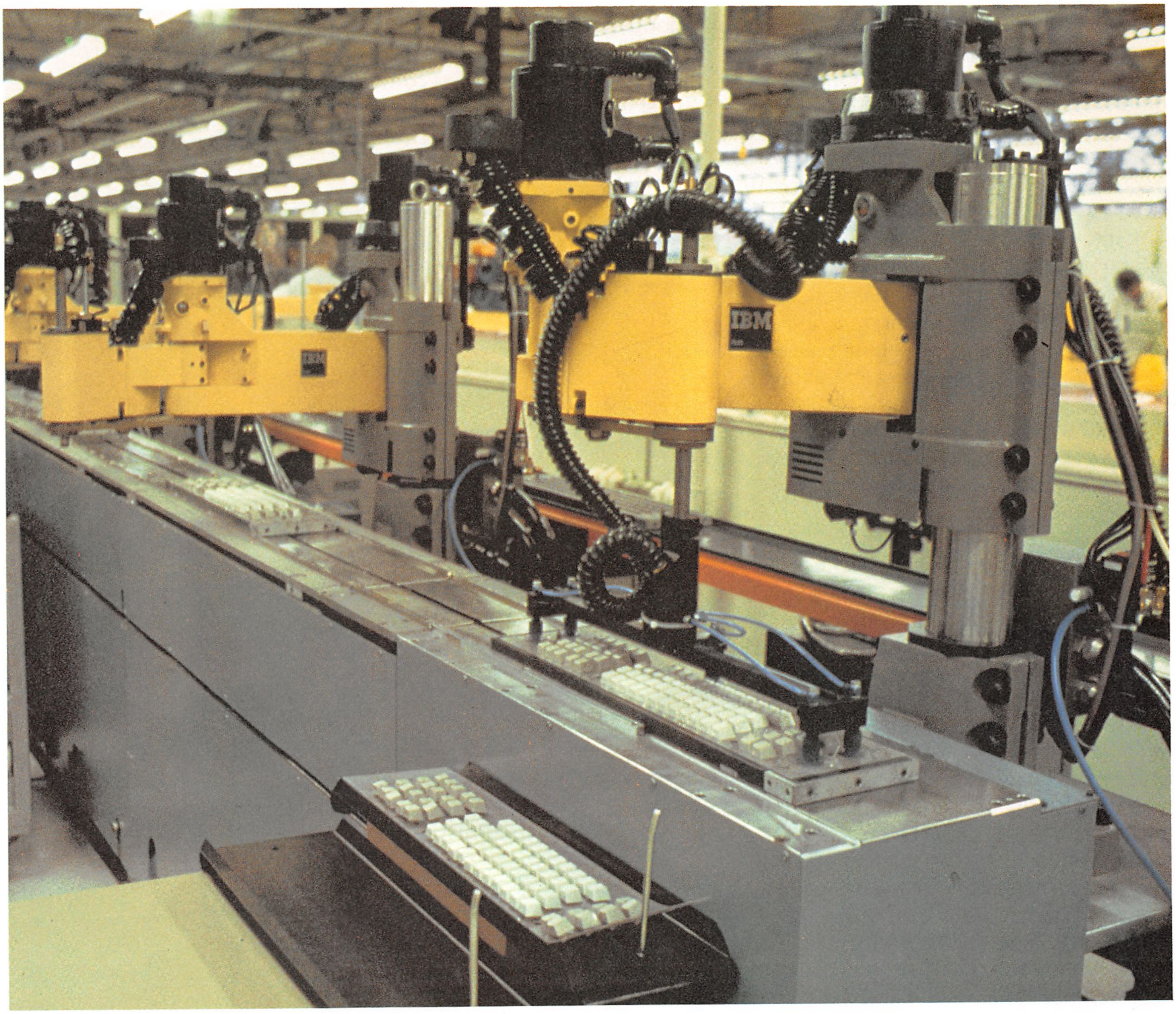
To work efficiently, all the technological advances in manufacturing industry have to work co-operatively. Designs created on CAD systems, for example, should be implemented on CAM machines – using materials created by a process controlled plant. Robot assembly completes the product, which is then automatically warehoused and distributed.

Simultaneously, management and sales force rely on computer generated data about products, sales and their financial position. Computers can also link directly with the production planners to provide the programs and operating sequence for the manufacturing plant.

Such a scenario requires reliable and cheap computers – which are available today. However, the barriers that are preventing our complete entry into the realm of the automatic factory are: our ability to plan and manage such complex computer systems; the balance between the large capital costs involved and the financial rewards; and our perception of the desirability of large scale automation, along with all the attendant social effects.

Paradoxically, it is society's demand for cheaper, more reliable and more complex products that is fuelling the drive towards the increased use of computers in design, management and manufacture.

Below: robots assembling IBM personal computers at Greenock, Scotland. (Photo: IBM).



Encoding satellite transmissions

COMMUNICATIONS

Multiple access

Multiple access in satellite communications is the term used to describe techniques for allowing a single satellite to be used simultaneously by several earth stations, whilst at the same time minimising interference between transmissions. A variety of encoding techniques are used which, at the receiving station, allow the receiving equipment to efficiently select only the signal meant for it, ignoring all others.

In general, multiple access methods involve the separation of signals in time, by frequency, by polarisation, or in space. We have just seen how signals can be separated in space, either by separating the orbital positions of the satellites themselves, or by the use of narrow beam earth aerials to focus closely on a single satellite. *Figure 1* illustrates how a satellite's spot beam can also provide separation in space.

We will consider signal separation by means of polarisation later in the chapter, and will now look at separation in frequency and time.

Frequency division multiple access
The separation of signals in the frequency domain is termed **frequency division multiple access** (FDMA), and it is based on the frequency division multiplexing techniques discussed in *Communications 3*.

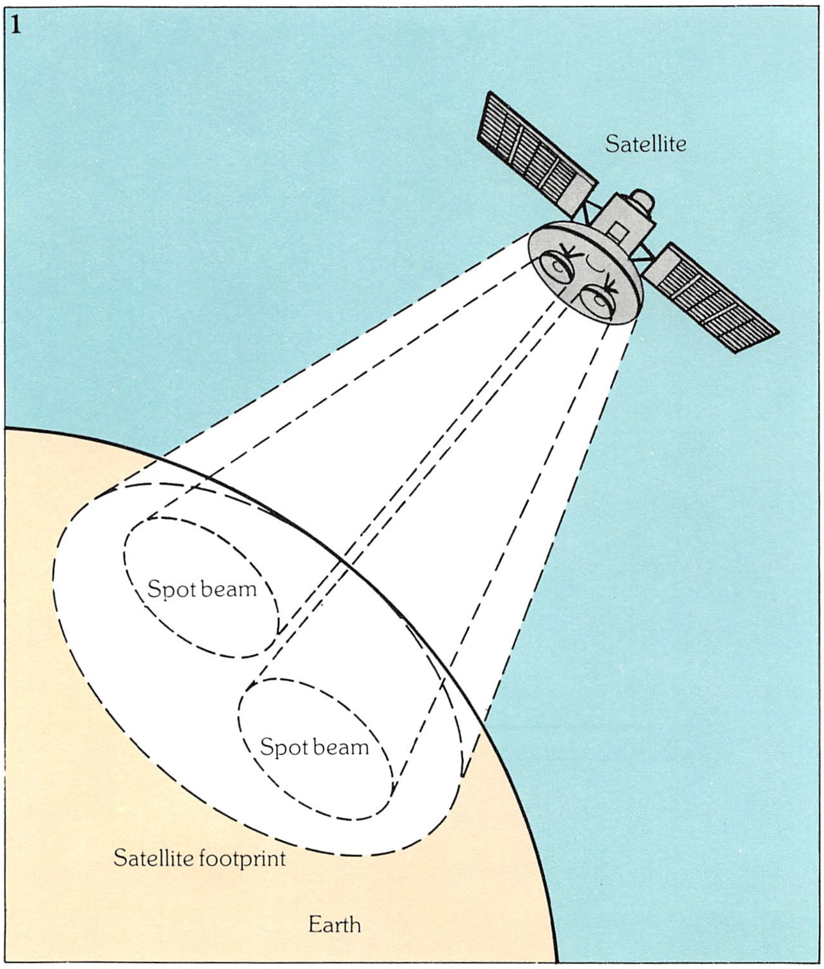
When signals are frequency division multiplexed, each signal source is assigned a range of frequencies adequate for the rate at which the user is transmitting information. Each signal is **translated** to a higher frequency by modulating a carrier frequency at the centre of the assigned bandwidth. Thus, each user's signal is translated to a new set of frequencies, with each separated so that there is no overlap which would cause mutual interference.

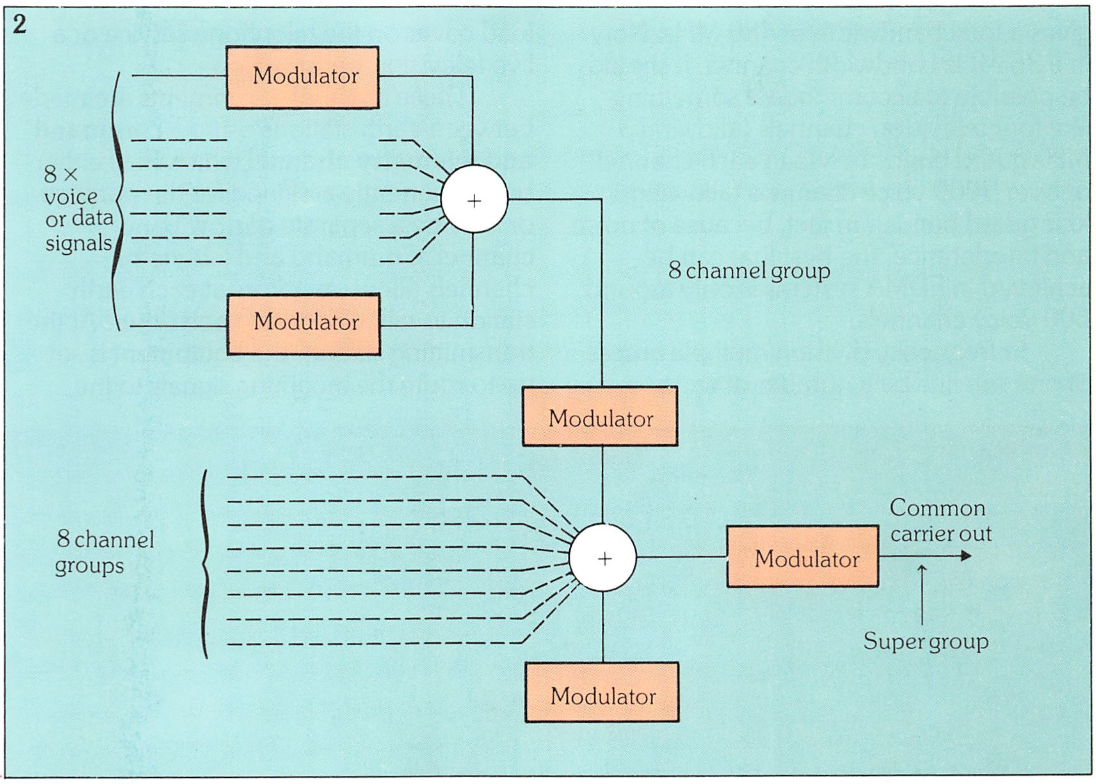
This group of multiplexed signals may then be used to modulate a common carrier (which may be transmitted by cable or radio) or it may be combined with another group of signals and further translated to form a super group.

Figure 2 illustrates how, by successive translation and combination, a large number of voice or data signals can be modulated onto a single carrier. Data signals, which in this case mean telex/teletype signals as well as computer data, must first be translated into the audio frequency range by frequency modulation methods.

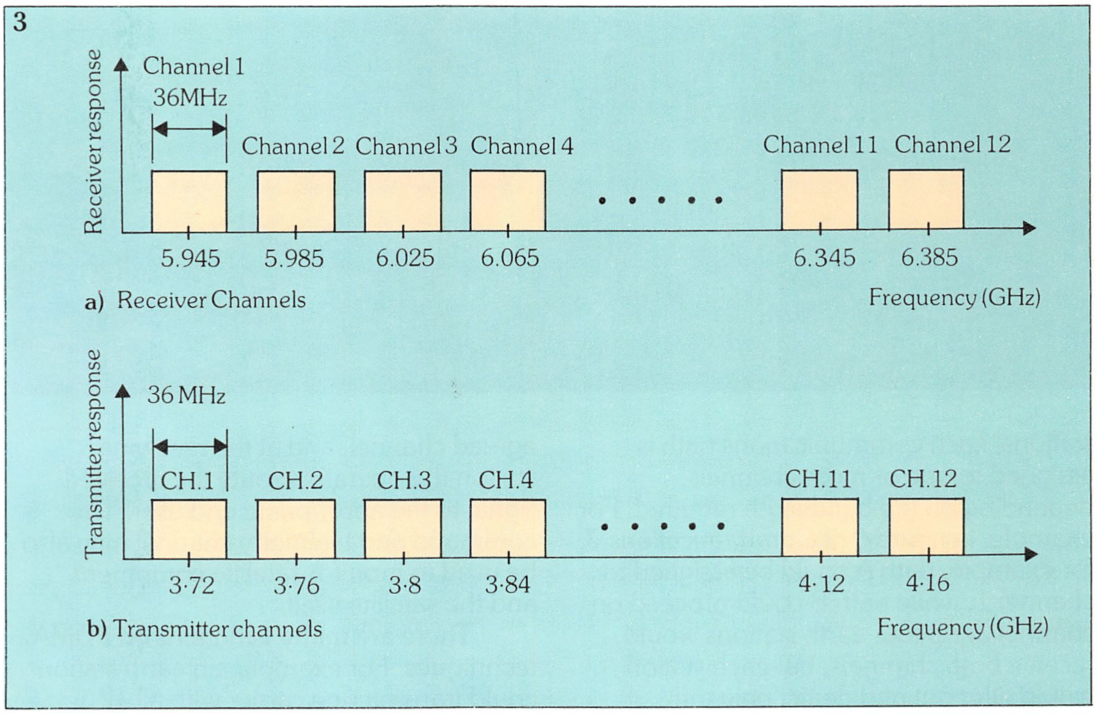
1. Satellite spot beams provide signal separation in space.

1





2. Successive translation and combination modulates a large number of voice or data signals onto a single carrier.



3. Frequency assignments for a typical satellite system.

From the discussion of frequency division multiplexing, we know that the limit to the number of voice/data channels that can be combined is the bandwidth of the communications link. However, much larger bandwidths are available in satellite communications systems, and so the process of translation and combination can be carried several stages further.

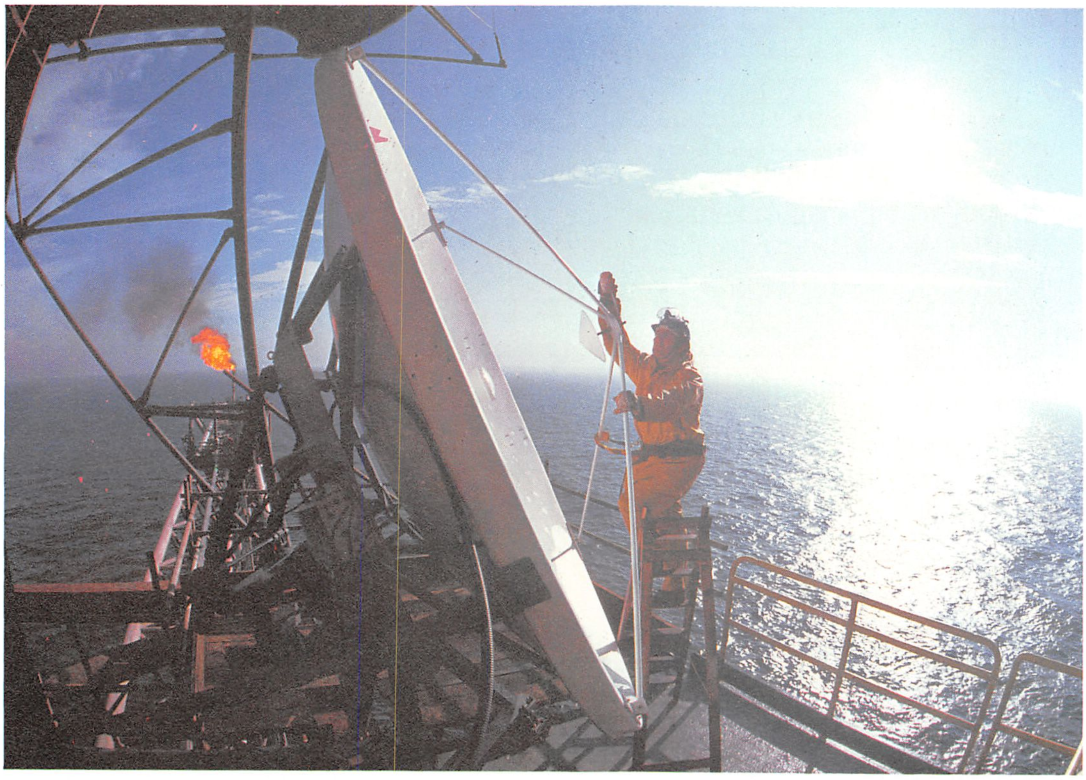
Figure 3 illustrates the frequency assignments for a typical satellite system. The system carries 12 channels, each 36 MHz wide and separated from adjacent channels by a guard band of 4 MHz. This

gives a total bandwidth of 480 MHz. Now in a 36 MHz bandwidth channel, it should be possible to accommodate something like four television channels (allowing 3 MHz guard bands between each channel) or over 9000 voice channels (allowing 1 kHz guard bands). In fact, because of noise and interference, the best that can be achieved in FDMA systems is only around 600 voice channels.

In frequency division multiple access of one satellite by a number of earth stations, each communications path is assigned to one or more channels, depending on the bandwidth required. For example, in *figure 6* of *Communications 7*, for example, path A could be assigned to channel 1, while path B could proceed on channel 2. All four earth stations would receive both channels, but each station would filter out and detect only the channel designated for it. Thus with FDMA, these four earth stations can communicate via a single satellite without overwhelming interference because each is separated from the others by frequency.

These channel assignments can either be permanent, as on the busy transatlantic telephone service, or temporary for peak load cover on the telephone service or a live television transmission.

These channel assignments are made between earth stations using a **command and telemetry channel** which may either be permanently assigned for the purpose or may be a separate narrow band channel. Command and telemetry channels allow operators at each earth station to talk directly to each other. At the transmitting station, the equipment is set up to route the incoming signals to the agreed channel, and at the receiving station the operators route the received signal to the appropriate end user. The command and telemetry channel may also be used to monitor satellite equipment, and the satellite itself.



Left: one of the many small dish earth aerials leased to business by British Telecom as part of its SatStream service. This one is on a North Sea oil rig. (Photo: British Telecom).

There are many variations on FDMA techniques. For example, an earth station could transmit one carrier with all 12 channels multiplexed onto it, with each destined for different receiving earth stations. Alternatively, each transmitting station could send a separate 36 MHz channel carrier for each designated earth receiving station.

A third alternative, known as **single channel per carrier** (SCPC), is to transmit

a separate carrier for each voice/data signal. In this case, each carrier is a frequency within one of the 36 MHz wide channels, but has a normal voice or data signal bandwidth. The separate carriers are then combined to form a signal with a 36 MHz bandwidth, rather than being multiplexed onto one carrier frequency, also with a 36 MHz bandwidth, as described earlier. This method enables greater efficiency in bandwidth use, particularly when used in the **SPADE** or **s**ingle channel per carrier, **p**ulse code modulation, multiple **a**ccess, **d**emand assigned equipment formats, as used on INTELSAT IVA.

The SPADE system utilises a single carrier located within a 36 MHz channel for each digitised voice or data signal from an end user, rather than multiplexing them together on a common carrier. One carrier within the 36 MHz channel is designated as a full time control channel and is used to assign a carrier to a voice or data channel, as required. These demand assignments can be carried out automatically, carriers being assigned and released to the pool on a continuous basis.

Using SPADE, it is possible to provide 800 voice/data channels (400 transmit/receive pairs) in each 36 MHz channel. This compares favourably with the 600 possible with conventional frequency division multiplex methods, simply because the available bandwidth is used more efficiently.

Another advantage of SCPC systems is an automatic keying system – this ensures that the carrier is transmitted only when there is a voice/data signal present, and results in considerable gains in transmitter power efficiency, as well as reducing interference between channels.

Analogue *vs* digital

Although, in principle, any FDMA system could carry analogue voice signals, digital signals are usually used as they present a number of advantages.

Historically, only analogue signalling was used, in satellite as in other forms of radio communications, and while the advantages of digital methods were realised at an early stage, they could not be put into operation until suitable equipment had been designed, developed and placed in orbit.

The greatest problem with analogue transmission is noise and, as mentioned earlier, signals buried deeply in noise or interference require far more bandwidth if they are to be received. The best capacity that FDMA systems can achieve, therefore, is about 16 channels per MHz, or 600 channels in a 36 MHz wide satellite channel.

Digital signals, on the other hand, are less susceptible to noise, and can be more easily processed using modern large scale integrated circuits. Although in principle they require more bandwidth per voice or data channel, digital signals actually permit a more efficient use of available bandwidth. In the SPADE system, each voice signal is digitised at a rate of 8000 samples per second, to a level of 8 bits per sample, i.e. a bit rate of 64 kbits s^{-1}.

After modulation, each digitised voice signal occupies a bandwidth of 38 kHz, and carriers, with one voice/data signal per carrier, are spaced at 45 kHz, thereby enabling a guard band of 7 kHz between them. Thus, each 36 MHz channel carries 800 individual voice/data channels, or 22 channels per MHz.

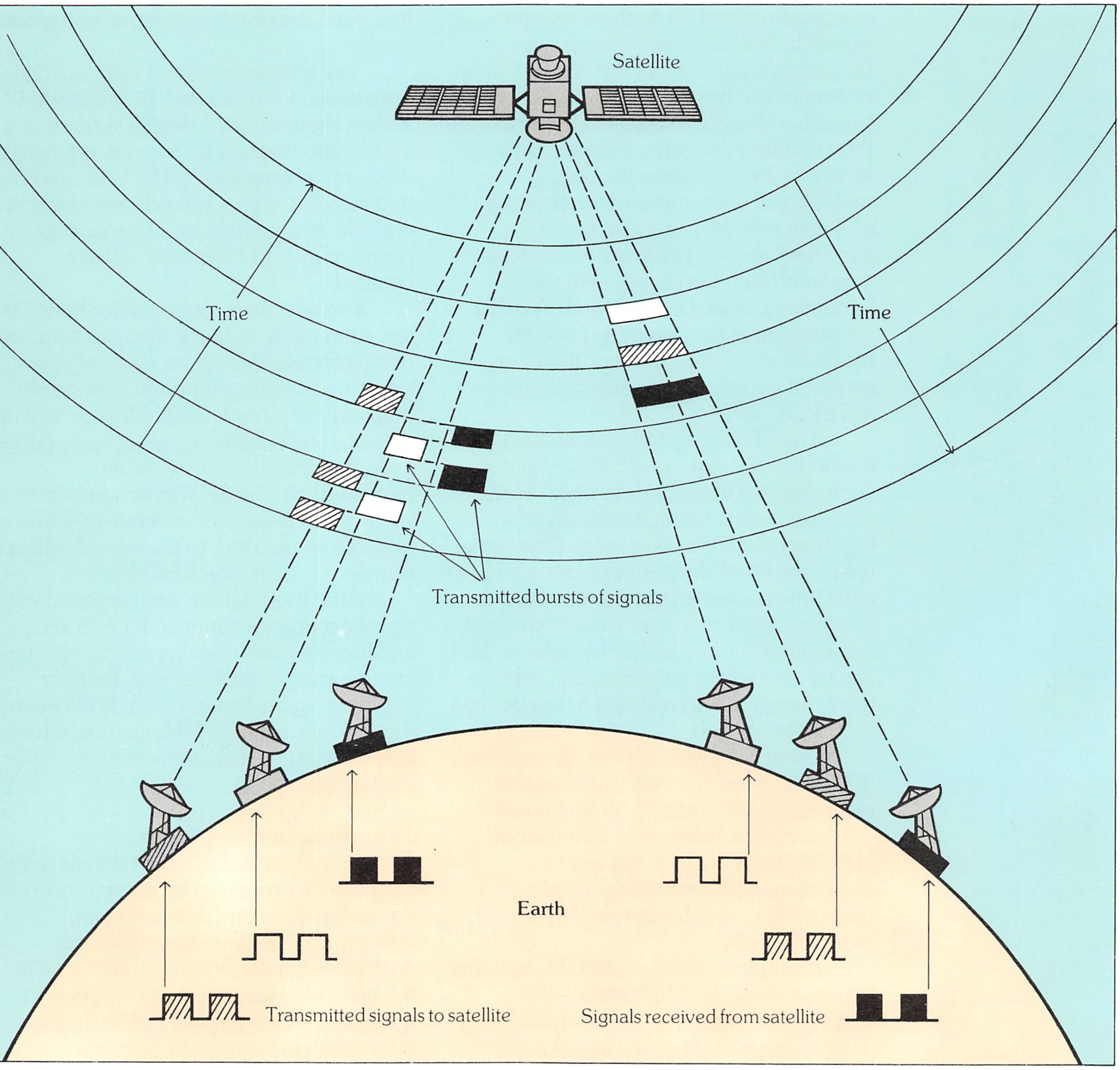
Time division multiple access

Another advantage of digital signals is that they can be processed by many powerful error correction and data encoding techniques that cannot be applied to analogue signals. One such technique is the basis for a second multiple access method which separates signals in time, rather than in space or in frequency.

Time division multiple access (TDMA) requires that every input to the system is either a digital signal (e.g. computer data or teletype/telex signals) or, if analogue (e.g. voice, video or some facsimile), that it be converted to digital form before being input to a TDMA system. This conversion process can be carried out at one of several points in the signal chain.

The basic principles behind time division multiplexing (TDM) have been discussed in *Communications 4* – successive samples from a digitised voice signal are assembled into a continuous

4



high speed bit stream, with samples from different sources interleaved in time. TDMA is a multiple access system for satellite communications using similar principles.

Each earth station in a TDMA network uses the same carrier frequency, but each transmits a **burst** of signal during a specific time period. Signals from a number of earth stations therefore arrive at the satellite at different times, and so cannot interfere with one another. The satellite transmits the time related signals to all receiving stations in the network, and each station selects only the signal which occurs at a designated time, ignoring all others, as shown in *figure 4*.

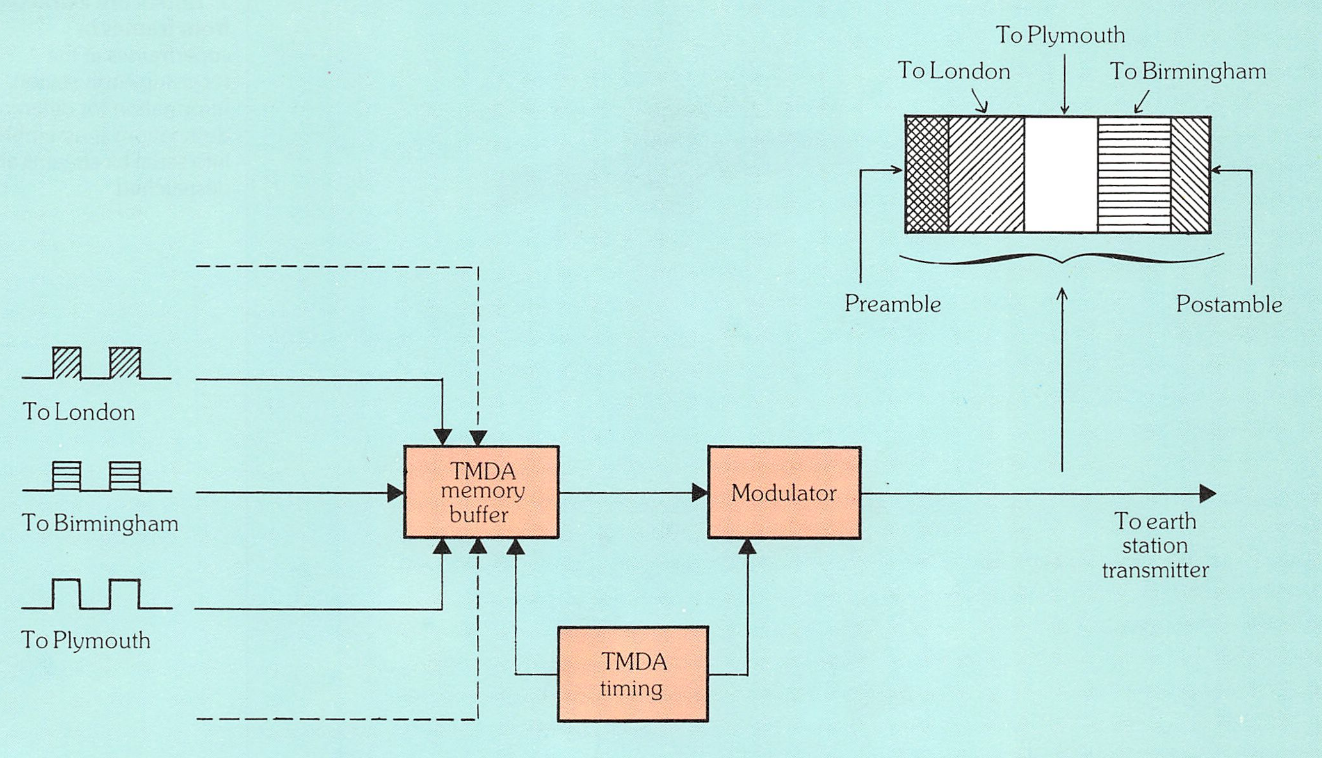
These time slots are arranged in advance of a transmission, so each receiving station knows that signals addressed to it will occur only at certain times. As with FDMA, the time slots can be arranged permanently or assigned as required. A common control channel between all earth stations is used to assign time slots and to release them to a common pool for re-use when a temporary transmission has been completed. As you

4. In a TDMA network, bursts of signals from a number of earth stations arrive at the satellite at different times.

5. Assembly of a burst of signals.

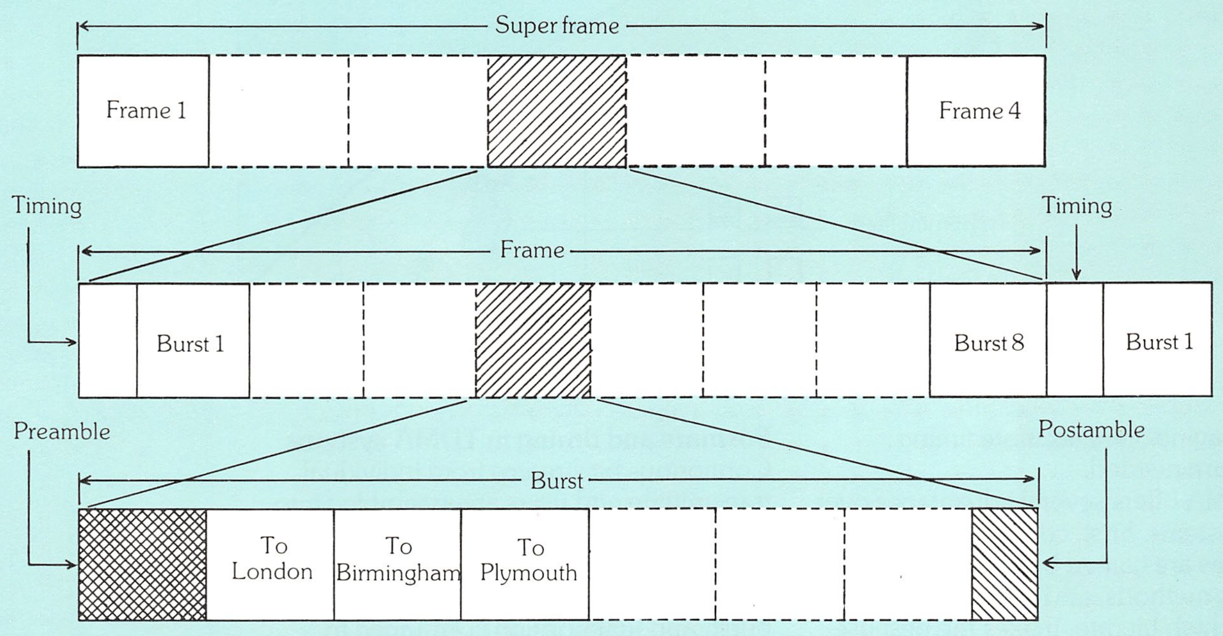
6. Eight bursts are linked into a frame. Frames can then be combined into superframes.

5

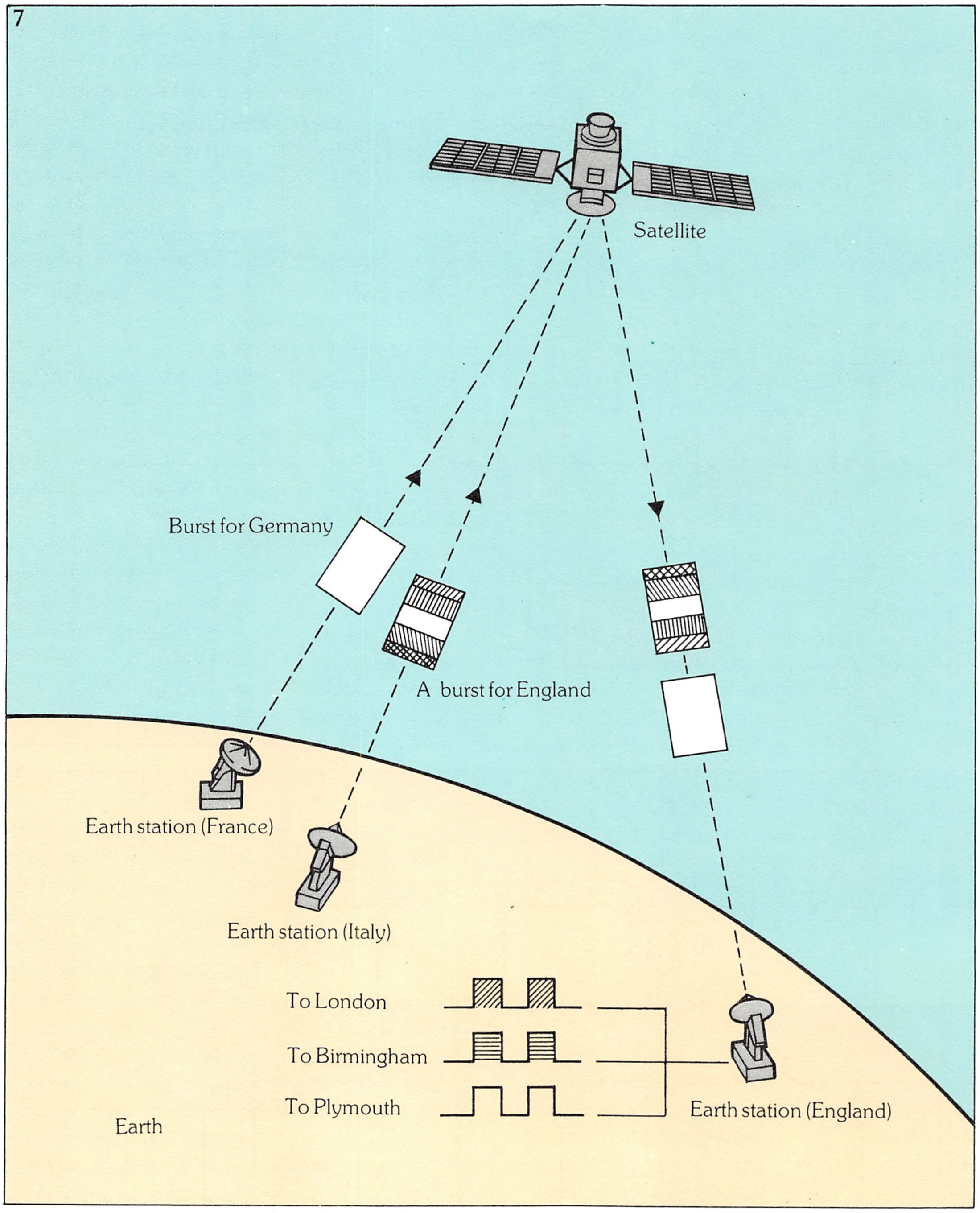


To Plymouth
To London
To Birmingham
Preamble
Postamble
To London
To Birmingham
To Plymouth
TMDA memory buffer
Modulator
TMDA timing
To earth station transmitter

6



Super frame
Frame 1
Frame 4
Timing
Timing
Frame
Burst 1
Burst 8
Burst 1
Preamble
Postamble
Burst
To London
To Birmingham
To Plymouth



7. Bursts are extracted from frames or superframes at the receiving earth station. Information for different destinations is assembled into serial bit streams and despatched.

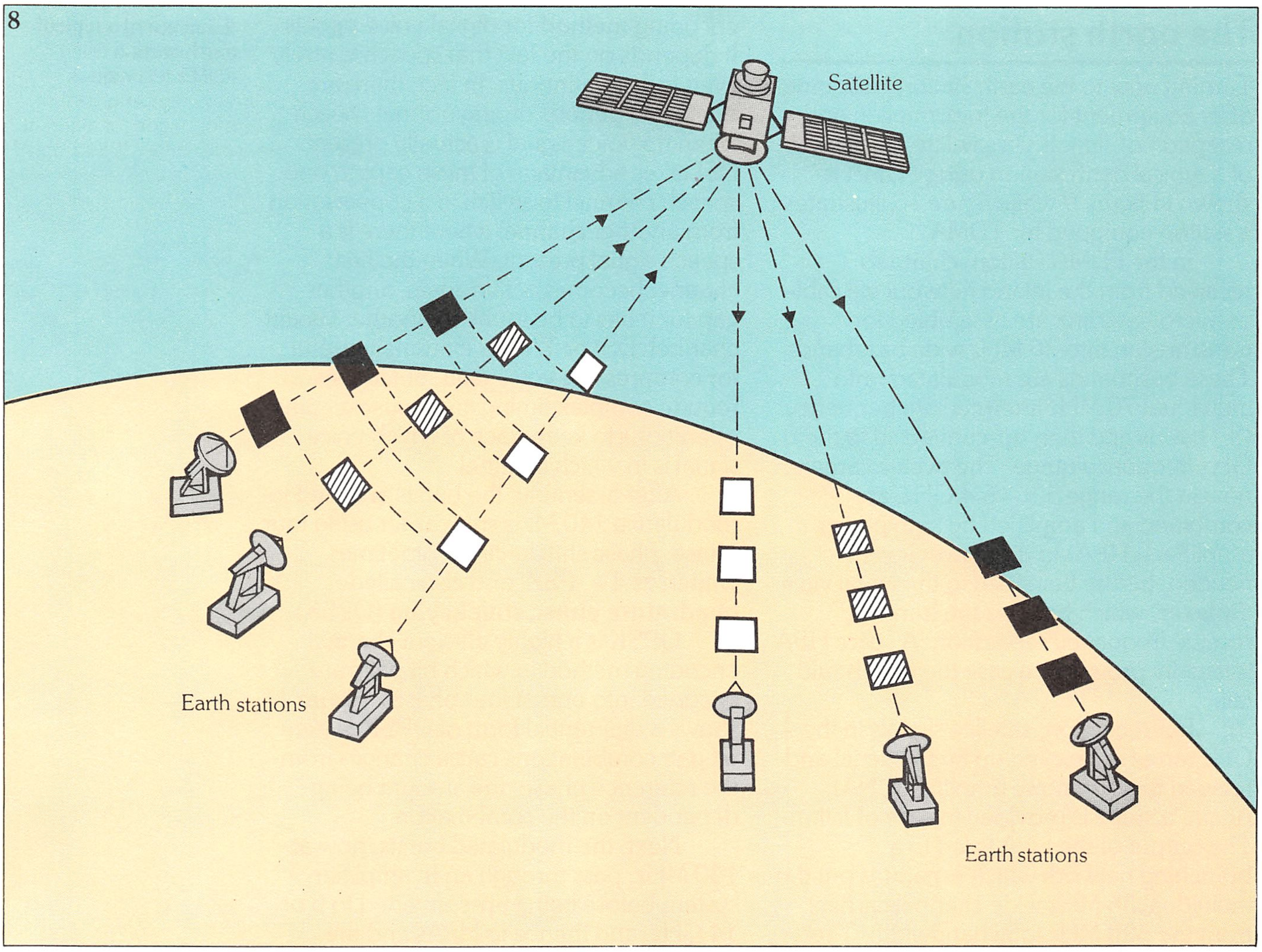
might imagine, very accurate timing systems are needed.

TDMA offers several advantages over FDMA systems. First, considerable efficiencies are gained by the use of digital switching methods, and TDM itself, with a naturally high bit rate, makes the best use of the available bandwidth. Second, since each station is transmitting on a single carrier frequency, transmitter efficiency is increased while the cost of building and operating an earth station is reduced.

Formats and timing in TDMA systems

Continuous bit streams from individual transmitting end users are assembled into bursts are shown in *figure 5*. Each bit stream, destined for a different receive end user, occupies a specific position within a burst, and that position is arranged to specify the address of the receiving end user. Each burst is preceded by a **preamble** – this indicates the start of a burst and contains information enabling the receiving equipment to maintain

8



8. SS-TDMA techniques enable the satellite to act as a switching centre in space.

synchronisation – and is followed by a **postamble** – this marks the end of the burst and is also used to maintain the frequency stability of the receiving circuits.

Bursts are then assembled into **frames** of eight bursts, with each frame preceded by essential timing information that holds the whole system together. Frames can also be assembled into **super frames**, as indicated in *figure 6*.

At the receiving earth station, super frames or frames allocated for that station are demodulated and the bursts extracted. The information intended for different end users is removed from the bursts, assembled as serial bit streams, and despatched via the inland telephone system to the appropriate address (*figure 7*).

Each earth station receives only those frames or super frames allocated to it, and each burst and each of the bit streams within a burst can be precisely identified by its time slot and can therefore be automatically routed to the correct destination.

A further development of TDMA techniques, **satellite switched TDMA** (SS-TDMA), promises even greater efficiency in using the available bandwidth. Using SS-TDMA techniques, a satellite not only acts as a 'transparent' repeater station but also as a switching centre in space as shown in *figure 8*.

Signals destined for a particular earth station may originate from three (or more) transmitting earth stations, as we have seen. With an on-board switching system, bursts or frames addressed to a certain earth station can be extracted from the time interleaved frame/burst sequence and gathered for transmission via highly efficient down-link spot beams aimed directly at the relevant station.

The earth station

Turning now to the earth station and some of its equipment for the transmission and reception of signals via satellite, the layout of a typical earth station using FDMA is shown in *figure 9* while *figure 10* illustrates a station equipped for TDMA.

In the FDMA station, channels received from the inland network via cable or microwave link are assembled for transmission into 36 MHz wide basebands. These basebands are modulated onto individual 70 MHz intermediate frequency (IF) bands and then **up-converted** to the 6 GHz frequency range. The twelve carriers, now in the range 5.9 – 6.4 GHz, are combined and amplified by high power amplifiers (HPA) to the power level needed, before being fed to the aerial via a **diplexer**, which handles the transmit/receive frequency separation. A spare HPA is usually provided in case the first should fail.

For reception, satellite signals in the 4 GHz range are picked up by the aerial and relayed to a low noise amplifier (LNA). Again, a spare is provided in case of failure. The output of the LNA is fed to a branching network which separates out the individual 36 MHz wide channel carriers from the 500 MHz total bandwidth. The carriers, in the range 3.72 – 4.2 GHz, are **down-converted** to the 70 MHz IF, and at this stage unwanted channels are removed from the system. After further IF amplification (where amplifiers are cheaper and more efficient than at microwave frequencies) the carriers are demodulated, demultiplexed and finally fed to the inland network for relay to telephones.

The set-up for a TDMA station, however, is quite different. Here, voice channels demultiplexed from the inland network link into individual bit streams and are fed to the TDMA terminal equipment. This converts the continuous bit streams into time interleaved bursts, as discussed earlier, and attaches preamble and postamble to each burst. At this stage, a technique known as **digital speech interpolation** (DSI) may be used to further increase the efficiency of the system.

DSI, also known as **time assigned speech interpolation** or TASI, is a data encoding method for digital voice signals. It depends on the fact that speech is rarely even truly continuous. In fact, there are significant periods during normal speech when no voice signal is actually present. DSI takes advantage of these gaps in one speech channel to switch in a conversation from another channel where there is a speech signal present. When the first channel becomes active again, another gap for it has to be found in another vacant channel. DSI or TASI is extremely useful for compressing voice data, but it does require complex timing and housekeeping operations to keep track of which voice signal is in which channel.

After assembly, the bursts are used to modulate a 140 MHz sub-carrier using four phase, phase shift keying, sometimes written as 4ϕ -PSK, and also called **quadrature phase shift keying** (QPSK).

QPSK is a highly efficient channel encoding method in which bit pairs are encoded into one of four phases. *Figure 11* shows a differential form of QPSK, where bit pair combinations cause changes from the reference phase, the amount being dependent on the combination.

Next, the modulated bursts, now at 140 MHz, pass through an IF amplifier system before being up-converted to 6 or 14 GHz, and thence to HPAs and the aerial. The receive chain functions in precisely the same manner, but in reverse.

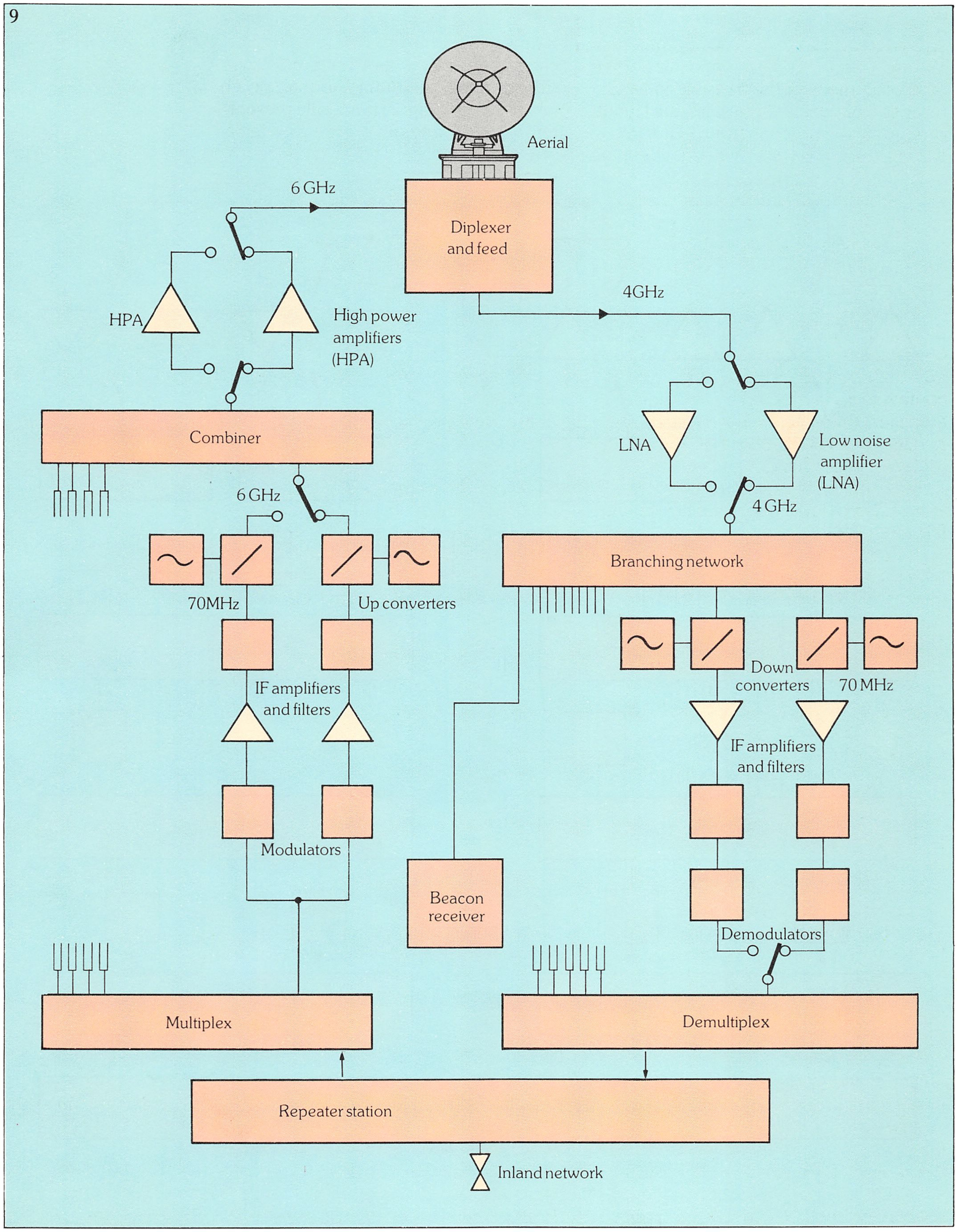
Klystrons and TWTs

Two types of amplifier are used as HPAs in satellite systems, the **klystron** and the **travelling wave tube** or TWT. Both are converters of DC energy to radio frequency energy in the microwave band, and both perform this by velocity modulation of a high speed electron beam (an electron beam is a direct current).

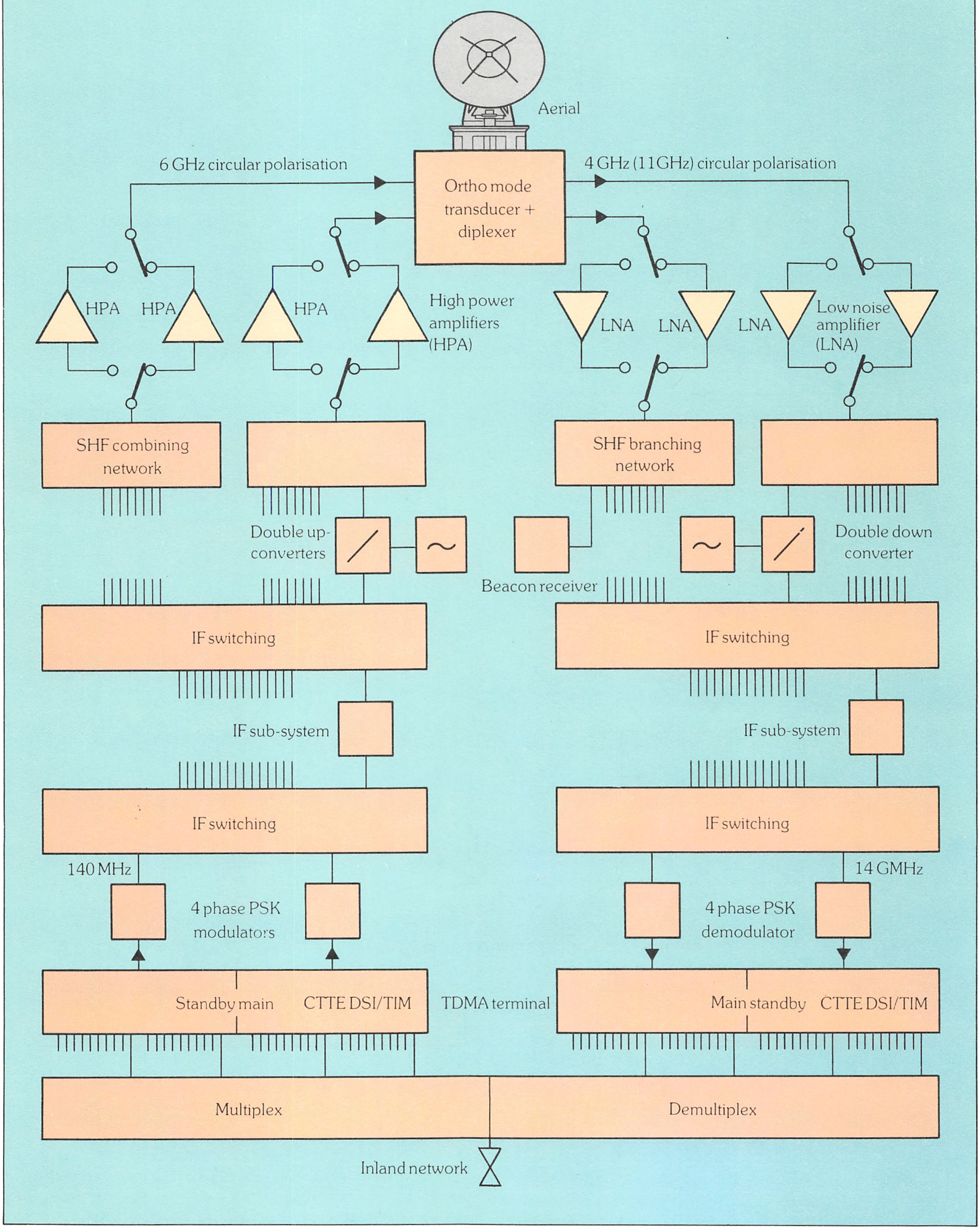
In the klystron (*figure 12*) an electron beam is formed at the cathode, focused, and accelerated to a high velocity which propels the electrons into the **drift space**. Surrounding the drift space are three **cavities**, and the first of these acts as an input transformer, coupling the RF energy input to the electron beam. This causes the velocity of some electrons to be reduced so that the beam becomes compressed at intervals corresponding to the frequency of

9. Layout of a typical earth station using FDMA techniques.

9


Aerial
6 GHz
Diplexer and feed
4GHz
HPA
High power amplifiers (HPA)
Combiner
6 GHz
LNA
Low noise amplifier (LNA)
4 GHz
Branching network
70MHz
Up converters
Down converters
70 MHz
IF amplifiers and filters
IF amplifiers and filters
Modulators
Beacon receiver
Demodulators
Multiplex
Demultiplex
Repeater station
Inland network

10


Aerial
6 GHz circular polarisation
Ortho mode transducer + diplexer
4 GHz (11GHz) circular polarisation
HPA
HPA
HPA
High power amplifiers (HPA)
LNA
LNA
LNA
Low noise amplifier (LNA)
SHF combining network
SHF branching network
Double up-converters
Beacon receiver
Double down converter
IF switching
IF switching
IF sub-system
IF sub-system
IF switching
IF switching
140 MHz
14 GMHz
4 phase PSK modulators
4 phase PSK demodulator
Standby main
CTTE DSI/TIM
TDMA terminal
Main standby
CTTE DSI/TIM
Multiplex
Demultiplex
Inland network

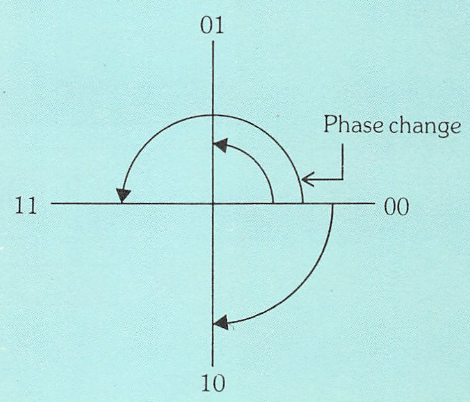
10. An earth station equipped for TDMA would look like this.

11. Quadrature phase shift keying (QPSK).

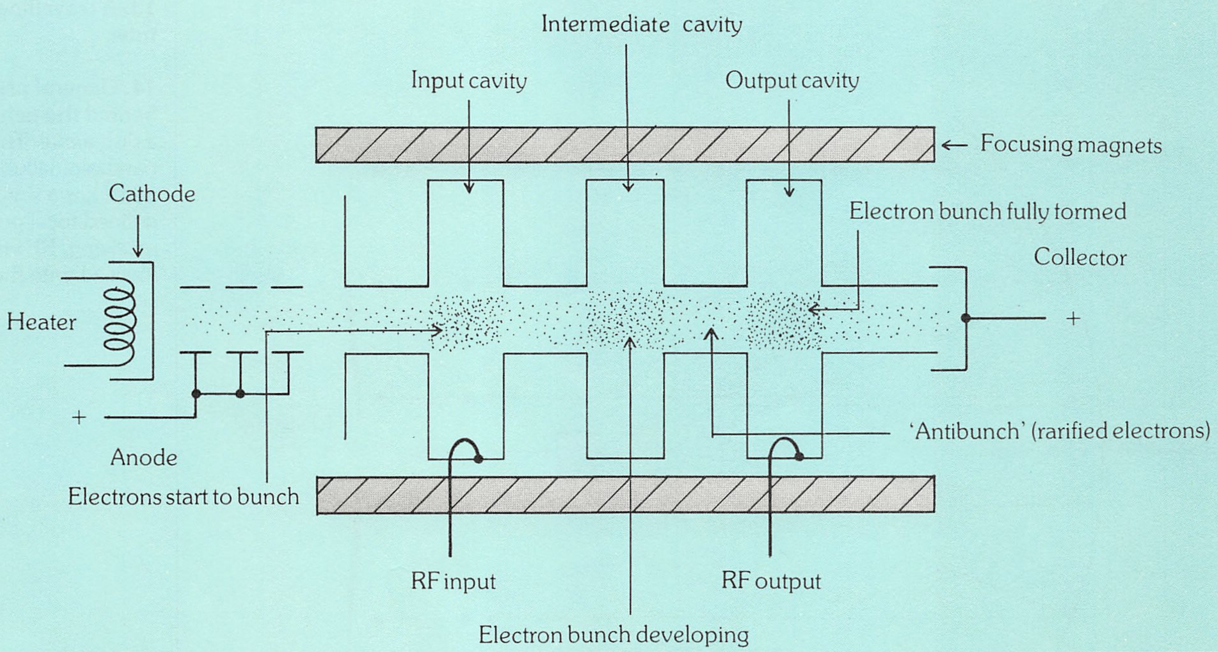
12. Action of a klystron.

11

Bit pair	Change
00	0°
01	90°
11	180°
10	270°



12

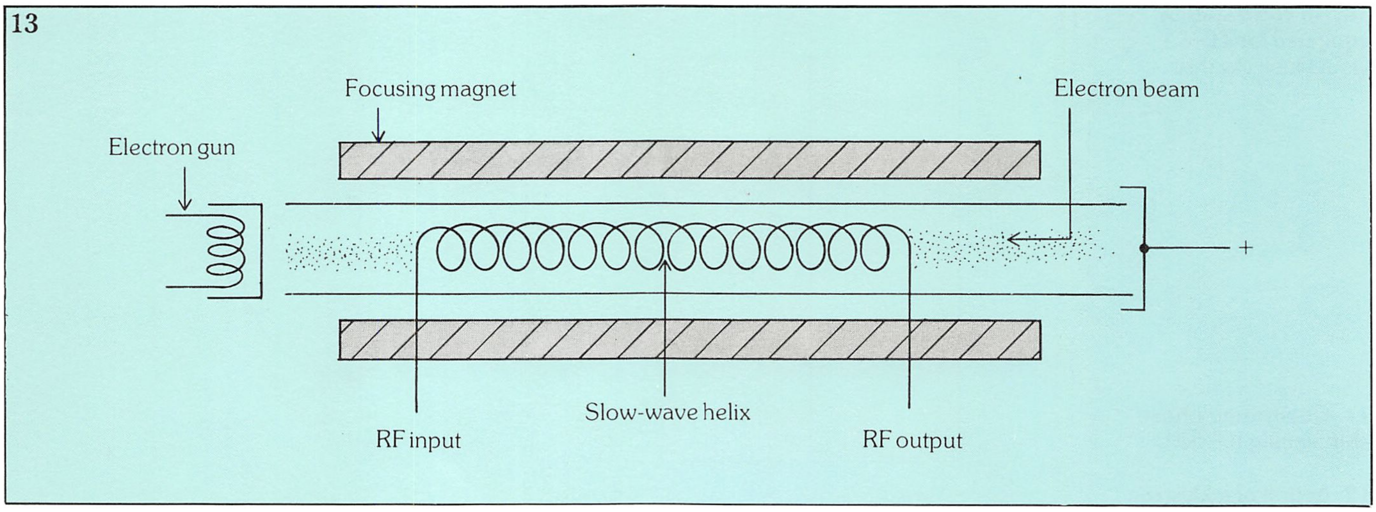


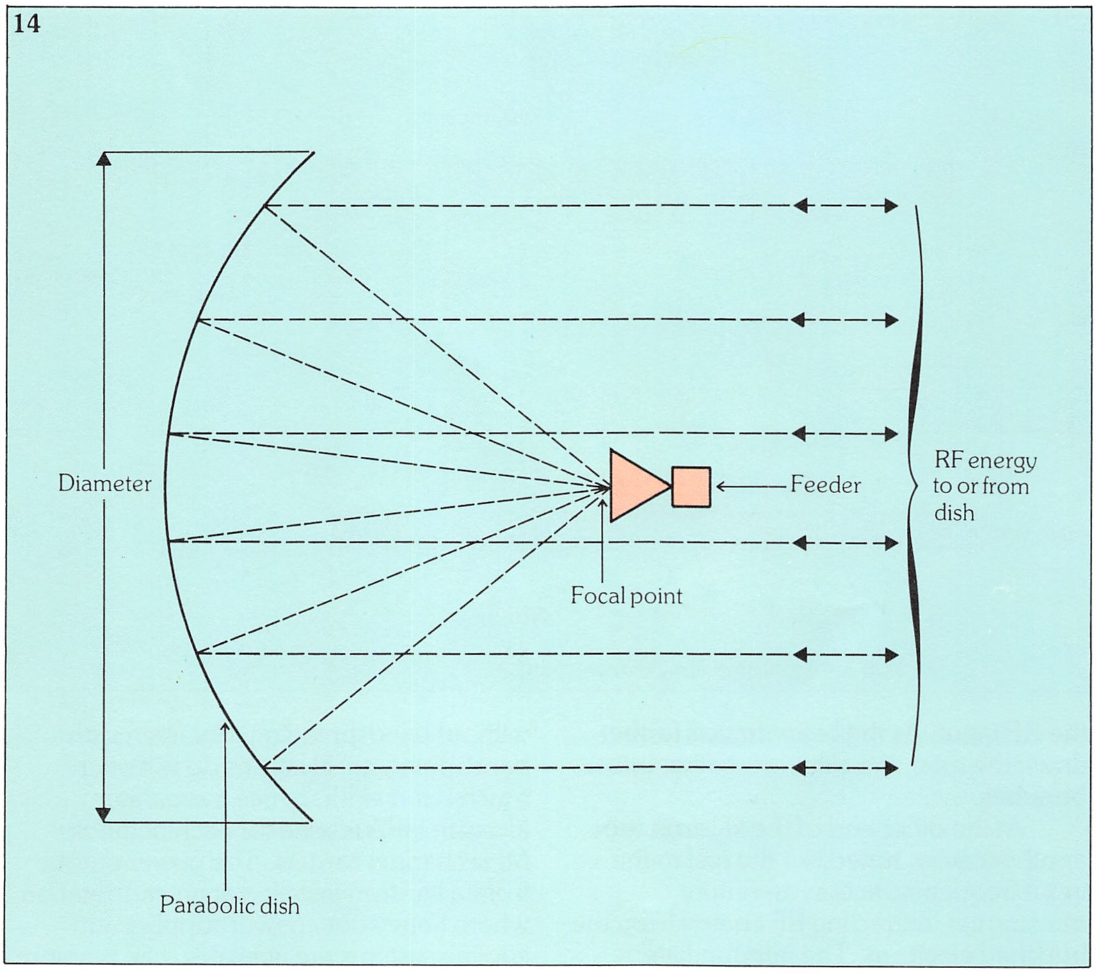
the RF input. As the beam travels further down the tube, the electrons gather into **bunches**.

At the other end of the klystron tube, another cavity, tuned like the first to the input frequency, acts as an output transformer, extracting RF energy from the bunched electrons. The intermediate cavity is not connected as an input or as an output, but assists in forming bunches, and increases the efficiency of the tube.

The klystron tube may actually be slightly de-tuned to broaden operating bandwidth – a process known as **bandspreading**. The overall gain of a three cavity klystron can be as much as 60 dB without bandspreading. But even *with* bandspreading, klystrons do not offer much bandwidth, hence a separate klystron HPA is used for each of the 36 MHz channel carriers. The power output from a klystron installed at an earth station, where heavy duty power supplies and cooling systems are available, can be more than 20 kW.

By contrast, TWT tubes operated as linear amplifiers have output power capabilities up to a kilowatt or so, but they do have the advantage of broadband amplification. Typically, bandwidths of between one and two octaves (represented by doubling a frequency) are possible,



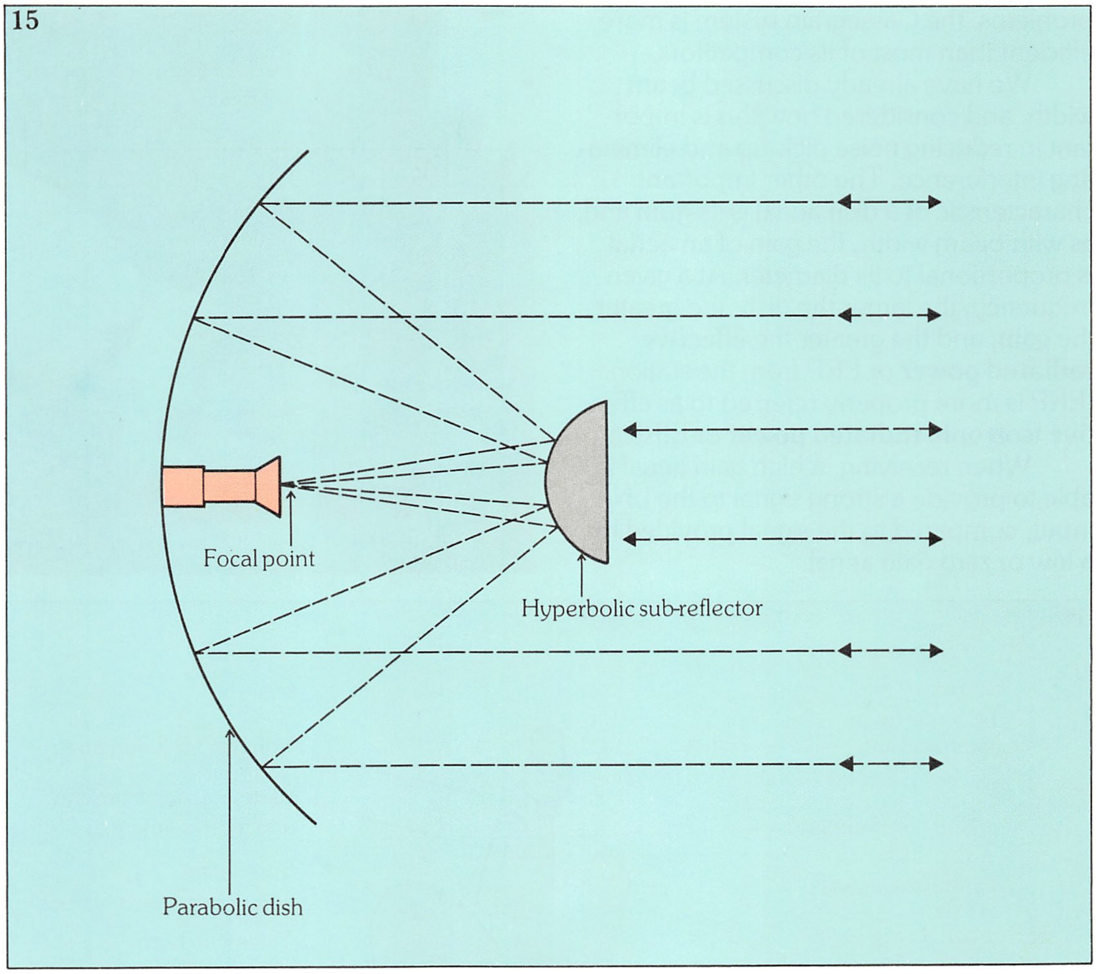


13. A travelling wave tube.

14. General principle behind the behaviour of a dish aerial. The paraboloidally shaped aerial has a single, defined focal point. If receiving, RF energy is focused onto the feeder.

depending on the power output required.

In a TWT, an electron gun fires a tightly focused electron beam along the length of a helix wound from round or flat wire. This coil is known as the **slow-wave structure**. The RF input is coupled to one end of the helix and the output is taken from the opposite end, as shown in *figure 13*. As in the klystron, the DC energy of the electron beam is converted to RF energy by velocity modulation of the beam. However, in the TWT, the beam interacts with the RF wave along the total length of the helix, rather than at cavity spaces.



15. Cassegrain reflector. This type of aerial, used by British Telecom, focuses microwave energy onto a feeder via a parabolic sub-reflector.

TWTs can also be used as low noise amplifiers (LNAs) in receiving equipment, but usually a class of amplifier known as a **parametric amplifier**, built around variable capacitance (varactor) diodes, is used.

Parametric amplifiers require liquid helium cooling systems for low noise operation, and are therefore slowly being replaced by semiconductor microwave amplifiers which use gallium arsenide FETs (GaAs FETs). These FETs are capable of providing gain in excess of 10 dB with a noise figure below 1 dB – all at 4 GHz, and without special cooling systems.

Aerials

The aerials used at earth stations are dish aerials – a section of a parabolic surface. They perform the same function in much the same way as a reflector in a torch.

In a transmitting aerial, RF energy from a **feeder** system is aimed at the dish, which focuses the radio waves into a parallel beam, just as a torch reflector focuses light from the incandescent bulb. If receiving, RF energy is collected by the dish and reflected as a tightly focused spot onto the feeder, which carries the signal to the LNA. *Figure 14* illustrates the general principle behind dish aerials.

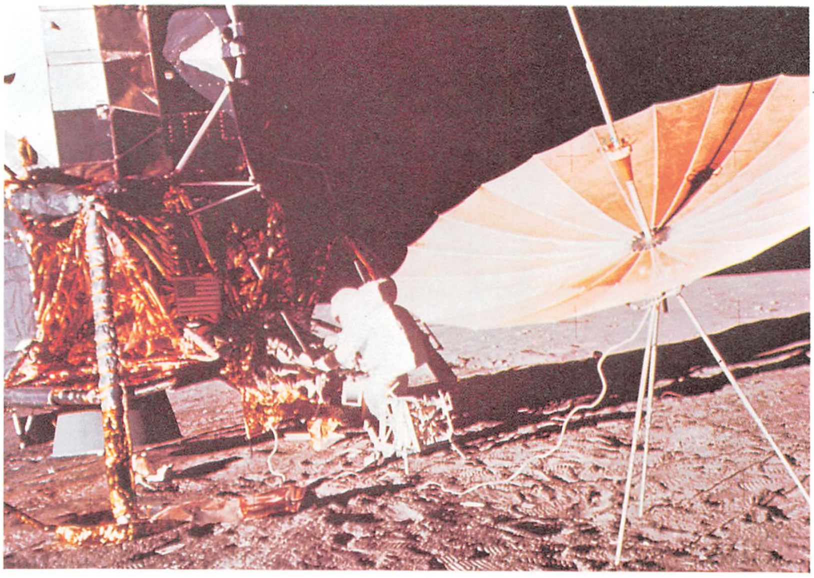
The particular type of system used by British Telecom at its Goonhilly Downs and Madley earth stations is shown in *figure 15*. These aerials are also known as **Cassegrain reflectors**, after the optical telescope system invented by the astronomer William Cassegrain. This system uses a parabolic sub-reflector to focus microwave energy onto the feeder.

You can see that more of the incident radio waves will be blocked from the aerial by the positioning of the sub-reflector, which is a disadvantage. However, the obstruction can be minimised by making the main dish very large by comparison to the sub-reflector, and this, of course, is possible at earth stations. Despite its

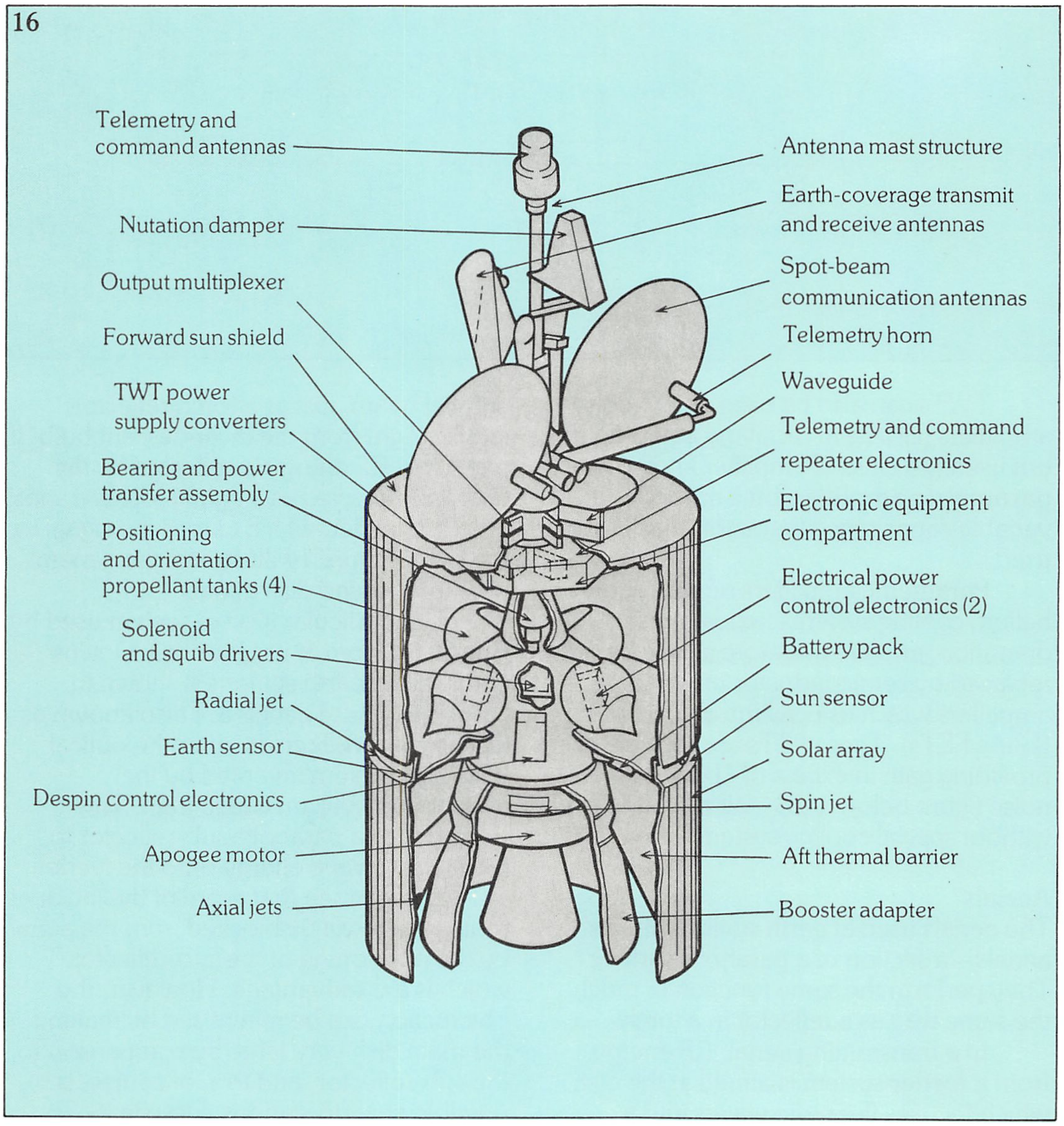
problems, the Cassegrain system is more efficient than most of its competitors.

We have already discussed beam width, and considered how this is important in reducing noise pick-up and eliminating interference. The other important characteristic of a dish aerial is its **gain** and, as with beam width, the gain of an aerial is proportional to its diameter. At a given frequency, the larger the dish, the greater the gain, and the greater the **effective radiated power** or ERP from the station. (ERP is more properly referred to as **effective isotropic radiated power** or EIRP.)

When receiving, a high gain aerial is able to provide a strong signal to the LNA input, compared to the signal provided by a low or zero gain aerial.



The Research House/NASA



Above: the Apollo 12 mission set up this antenna on the moon in 1966. The antenna transmits and receives RF energy in the S-band of microwave frequencies, which range from 1.55 to 5.20 GHz.

16. Cut-away drawing, showing the constituent parts of a satellite. This is based on INTELSAT IV.

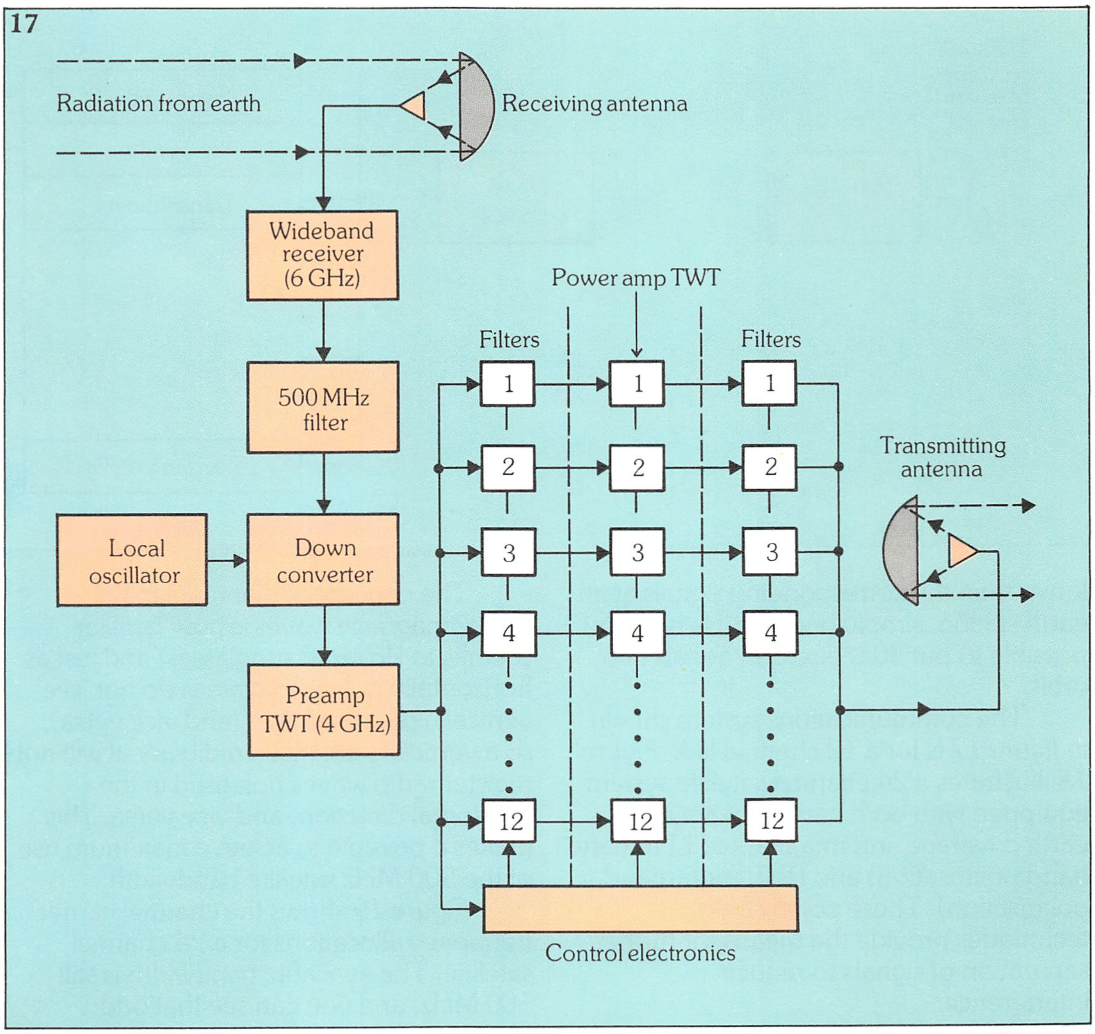
The satellite

The satellite shown in *figure 16* is based on the INTELSAT IV satellite, and a block diagram illustrating the communications equipment that such a satellite would carry is shown in *figure 17*. All electronic equipment is powered by solar energy, converted to electrical energy by solar panels.

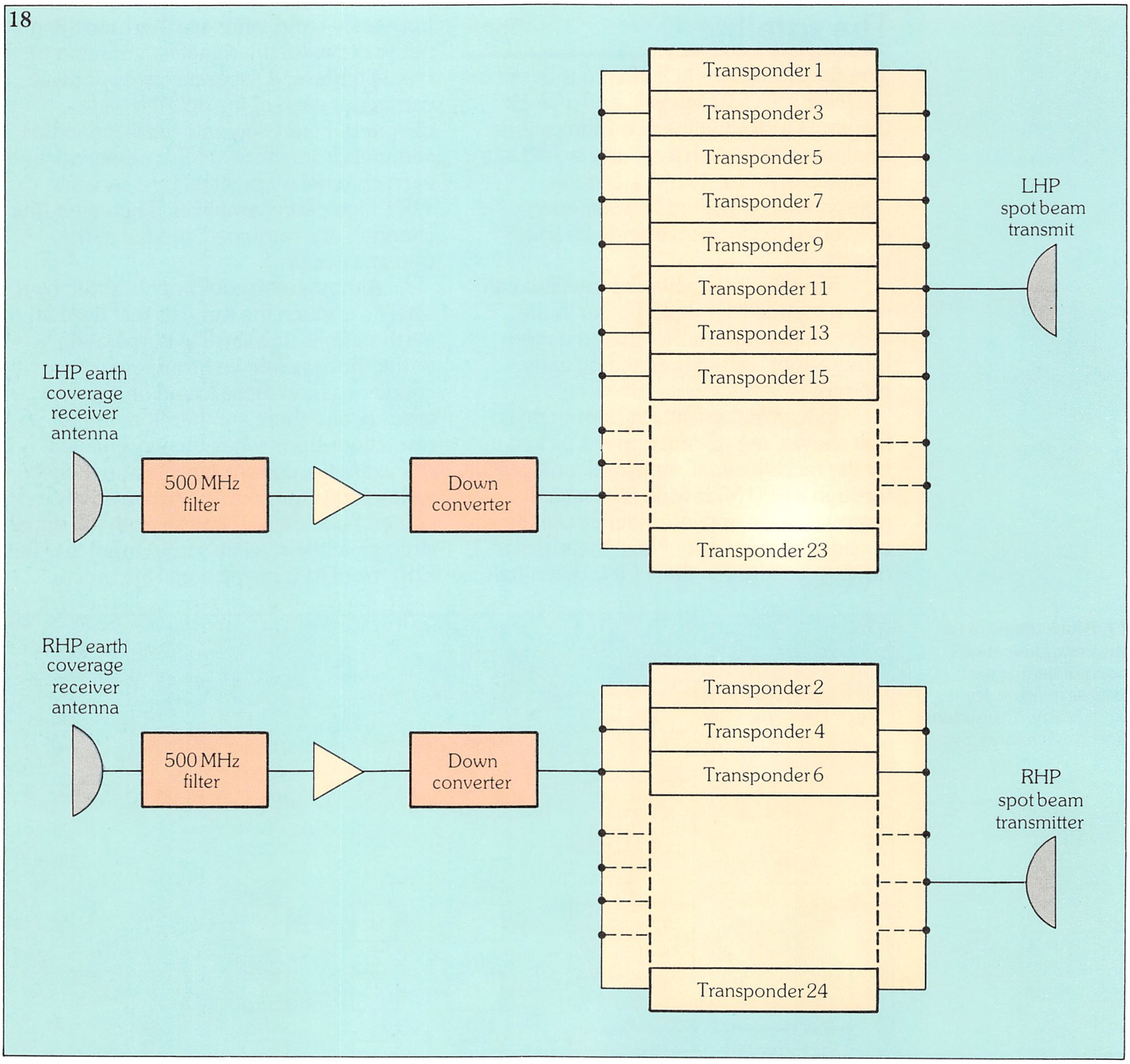
Modern solar powered satellites can now generate RF output power in the order of kilowatts, rather than the mere hundreds of watts generated by early models.

In a typical satellite system such as that shown, the up-link signal is picked up by the receive aerial, amplified, and passed through a 500 MHz wide filter which removes noise and interference signals outside the passband. Next, the signal is down-converted to the 4 GHz down-link frequency band, after which a branching network passes the signal to a number of **transponders**. A separate transponder is carried for each of the 36 MHz wide channels. Filters separate these individual channels from the 500 MHz passband, and each channel is amplified by a separate TWT high power amplifier. Thereafter, the channels are combined and fed to the transmit aerial.

Although the equipment carried by a satellite is much the same as that used at earth stations, the satellite is operating under more severe environmental conditions than those found on the ground, and there are therefore many restrictions upon what can be achieved. For example, current satellite and rocket technology limits the size of the aerials carried by a satellite, and so both the signal strength at the satellite's LNA input and the EIRP from its transmit aerial are much



17. Block diagram of the communications equipment that the satellite in *figure 16* would carry. This system is for a 12 channel link.



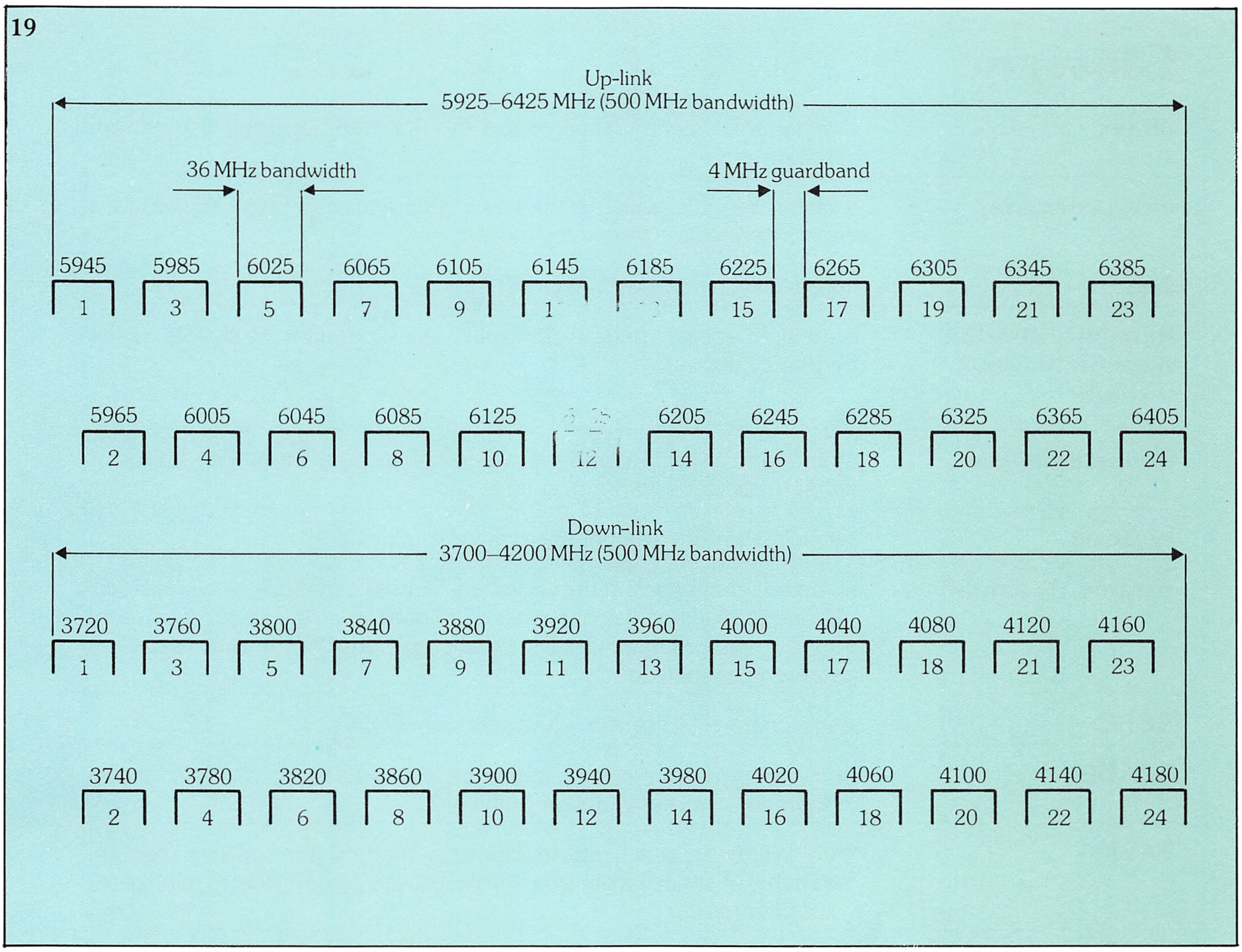
18. 24 channel satellite communications system.

lower than the corresponding signals at an earth station, simply because it is not (yet) possible to put 30 metre dish aerials into orbit.

The communications system shown in *figure 17* is for a 12 channel link. *Figure 18* illustrates a 24 channel satellite system equipped with both narrow beam and earth coverage antenna marked LHP (left-hand polarisation) and RHP (right-hand polarisation). These polarisation techniques provide the means for further separation of signals to reduce interference.

The concept of plane polarised electromagnetic waves is now familiar (thanks to Polaroid sunglasses) and just as horizontally polarised glasses do not 'see' vertically polarised light (and vice versa), so a vertically polarised radio aerial will not register radio waves polarised in the horizontal direction, and vice versa. This makes it possible to achieve maximum use of the 500 MHz satellite bandwidth.

Figure 19 shows the channel carrier frequency allocations for a 24 channel satellite. The available bandwidth is still 500 MHz, and you can see that odd



19. Channel carrier frequency allocations for a 24 channel satellite.

numbered channel bandwidths overlap with even numbered channels, with potentially disastrous results in terms of cross-channel interference. However, if the odd channels are transmitted and received with vertical polarisation while the even channels are horizontally polarised, they are effectively isolated from each other. Thus, polarised transmissions make it possible to use some frequencies more than once. In general, methods which allow this are called **frequency re-use** techniques.

Conclusion

Satellite communications has been one of the most extraordinary technological developments of the 20th century. From the germ of an idea proposed by Arthur C. Clarke in 1945, there is now a global network of satellites for international communications plus a growing number of satellites for domestic communications and Direct Broadcast Satellite television.

Yet there is always a demand for more channels, more links, and more satellites. The day is not far off when, to fill those needs, the vital equatorial orbit will be filled to capacity, with satellites spaced as closely as the aerial beam width and receiving technology will allow.

However, new methods for increasing capacity, such as TDMA and SS-TDMA, will add to the number of channels and paths available through a satellite, while inter-satellite links will, in one stride, greatly add to the efficiency with which those links are used. The day is also not far off when personal satellite dish aerials are installed on roof tops for quick and flexible communications to any point on the globe.

Glossary

diplex	simultaneous communications on two frequencies using a common aerial system
down-converter	a circuit module which translates high frequency radio signals to a lower intermediate frequency band
DSI	digital speech interpolation (see TASI)
frequency division multiple access (FDMA)	a multiple access method that relies on separation of signals in the frequency domain
frequency re-use	methods in satellite communications for using a frequency band for two simultaneous communications paths
klystron	a special vacuum tube for microwave amplification
parametric amplifier	an amplifier in which an input voltage causes variations in an electrical parameter, such as capacitance, in a linear manner, to provide gain. Variable capacitance diodes (varactors) are commonly used in parametric amplifiers
SCPC	single channel per carrier – a variation on FDMA
SPADE	single channel per carrier, pulse code modulation, multiple access equipment: another variation on FDMA
TASI	time assigned speech interpolation: a method for splicing digitised fragments of speech from one conversation into gaps found in another voice channel
time division multiple access (TDMA)	a multiple access method that separates signals in time
transponder	satellite communications equipment for down-converting the up-link signal to the frequency band used for the down-link
TWT	travelling wave tube: vacuum tube device for microwave amplification
up-converter	a circuit module which translates intermediate signals to a higher band

ELECTRICAL TECHNOLOGY

Impulses and network response

One of the most common problems which occurs in analogue networks is how to determine the output of electrical network when the input consists of some waveform such as speech or music signals. In the previous two *Basic Theory Refreshers* we have seen how such apparently formless signals can be broken into a number of sinusoidal signals whose frequencies cover a wide range.

Also in earlier *Basic Theory Refreshers* we have seen the way in which two-port networks can be analysed for sinusoidal input signals. This therefore provides one method of determining the output of a network to any arbitrary input. One limitation of this method, however, is that a Fourier analysis needs to be performed on the input waveform; the network at each frequency of the Fourier series must then be analysed, and finally, all these individual reponses must be added together to obtain the complete network response.

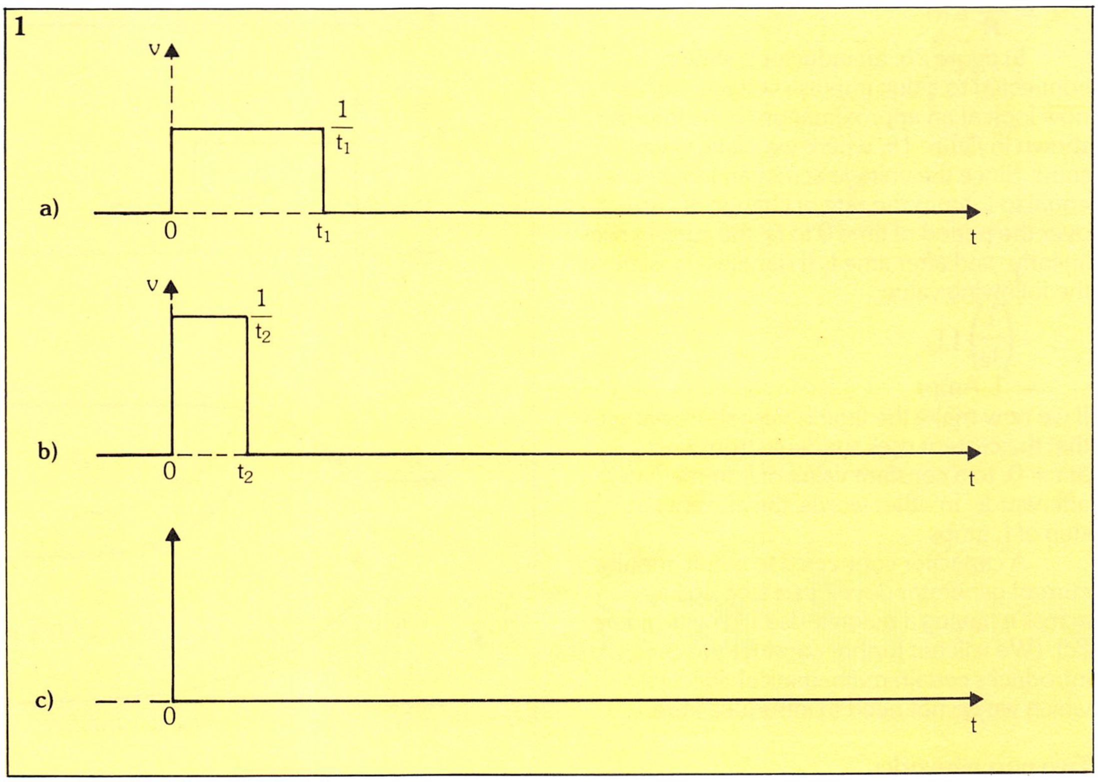
There is another way in which this result can be obtained, but first we must consider a new form of function, the **impulse function**.

The impulse function

Let us consider the pulse of voltage shown in *figure 1a*, having a rectangular shape, duration t_1 and height $1/t_1$; the area under this pulse will be $t_1 \times 1/t_1 = 1$. If the pulse duration is now reduced to t_2 and, at the same time, its amplitude is increased to $1/t_2$, then the area under the pulse will still be 1 as shown in *figure 1b*.

If we continue to reduce the pulse duration while at the same time increasing the amplitude, so that the area under the pulse remains at 1, we arrive at the situation where we have a pulse of infinitely short duration and infinite height, but which still has an area of unity. This, of course, is a very artificial idea and in practice it is always thought of as being a very short pulse, with a very large height. Such an idealised pulse is termed a **unit impulse** voltage and is usually written as δ (t) (it is sometimes known as a **Dirac function**). The unit impulse is drawn diagrammatically as an arrow pointing upwards as shown in *figure 1c*.

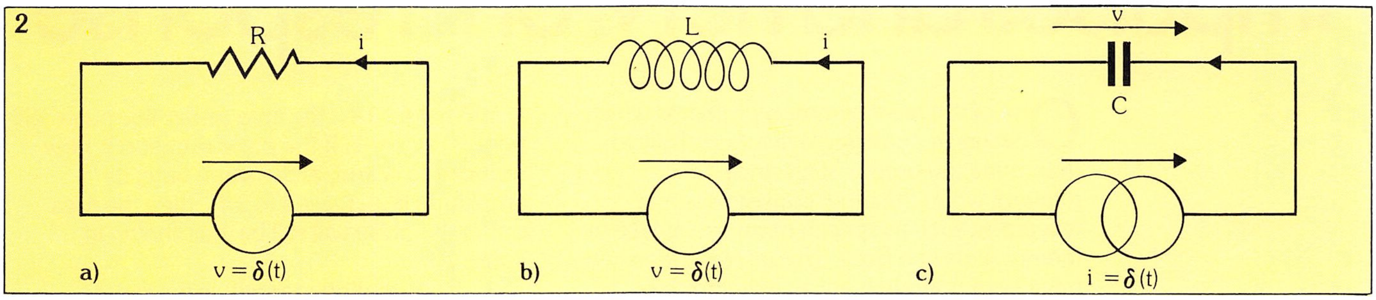
It is interesting to note that the Fourier integral of a unit impulse consists of a spectrum in which the amplitude of all the elements is uniform, from very low frequencies up to very high frequencies.



1. A pulse of infinitely short duration and infinite height, with an area of unity, is known as a unit impulse.

Response of single circuit elements

We will now consider the behaviour of resistors, capacitors and inductors when they



are supplied by a voltage or current which is a unit impulse.

A resistor connected to a unit impulse voltage generator is shown in *figure 2a*. Since the current flowing in a resistor is always given by Ohm's law as being proportional to the voltage, the current is also an impulse given by:

$$i = \frac{1}{R}\,\delta(t)$$

In *figure 2b*, an inductor is shown connected to a unit impulse voltage. Let us now look at an approximation to the impulse shown in *figure 1b*, where we make t_2 very short. Since the voltage across an inductor is equal to L times the rate of change of current, over the period of time 0 to t_2, the current rises linearly, and after time t_2 it remains constant at the following value:

$$i = \left(\frac{1}{t_2}\right) Lt_2$$
$$= L \text{ Amps}$$

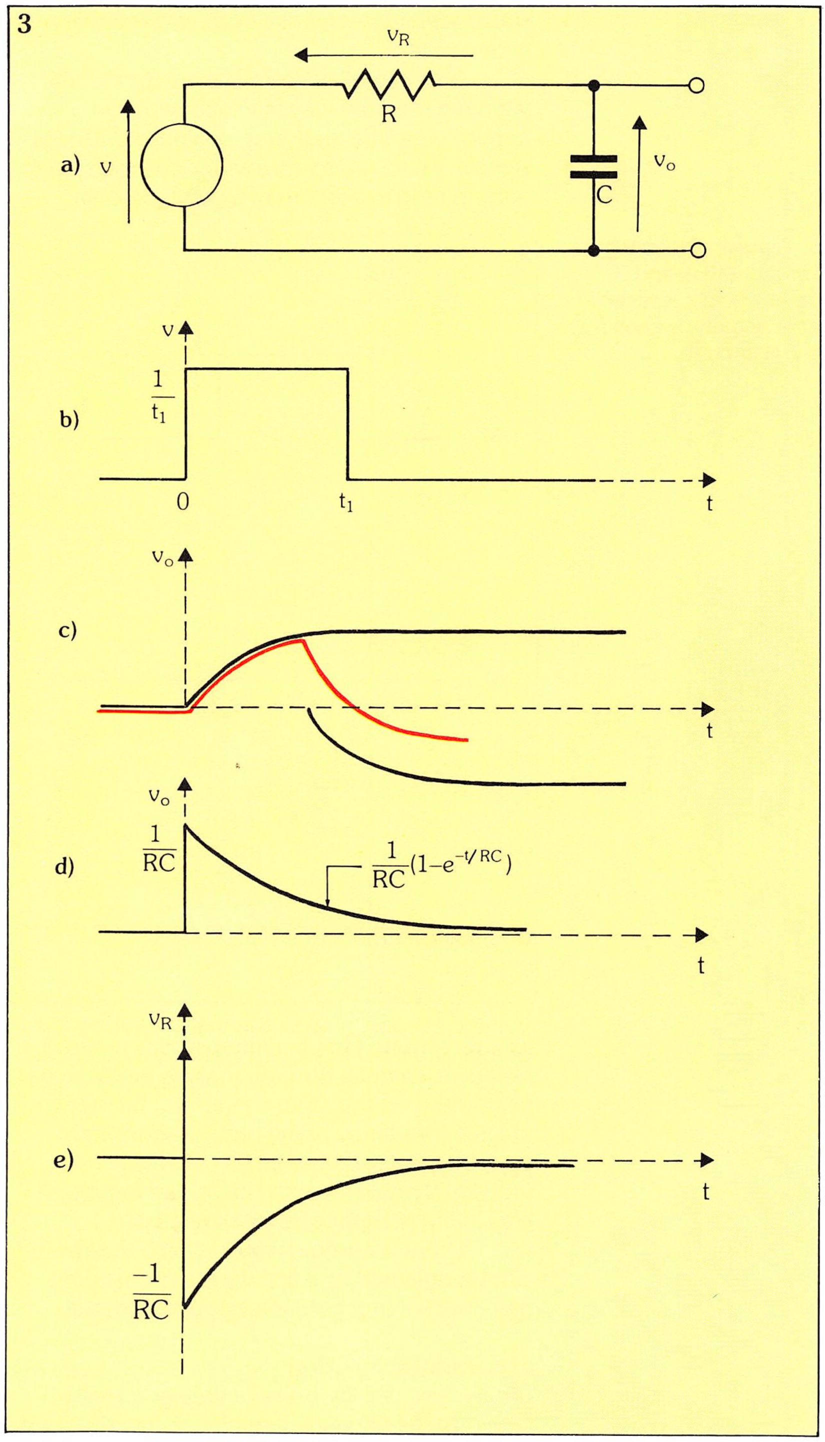
If we now make the time t_2 very short, we see that the current rises suddenly from zero, at $t = 0$, to a constant value of L immediately afterwards. In other words, the current rise is a step of L amps.

A capacitor connected to a unit impulse current generator develops a step voltage across it having a magnitude of C volts (*figure 2c*). (We will not further consider the circuit as it introduces certain mathematical difficulties which we do not need to introduce here.)

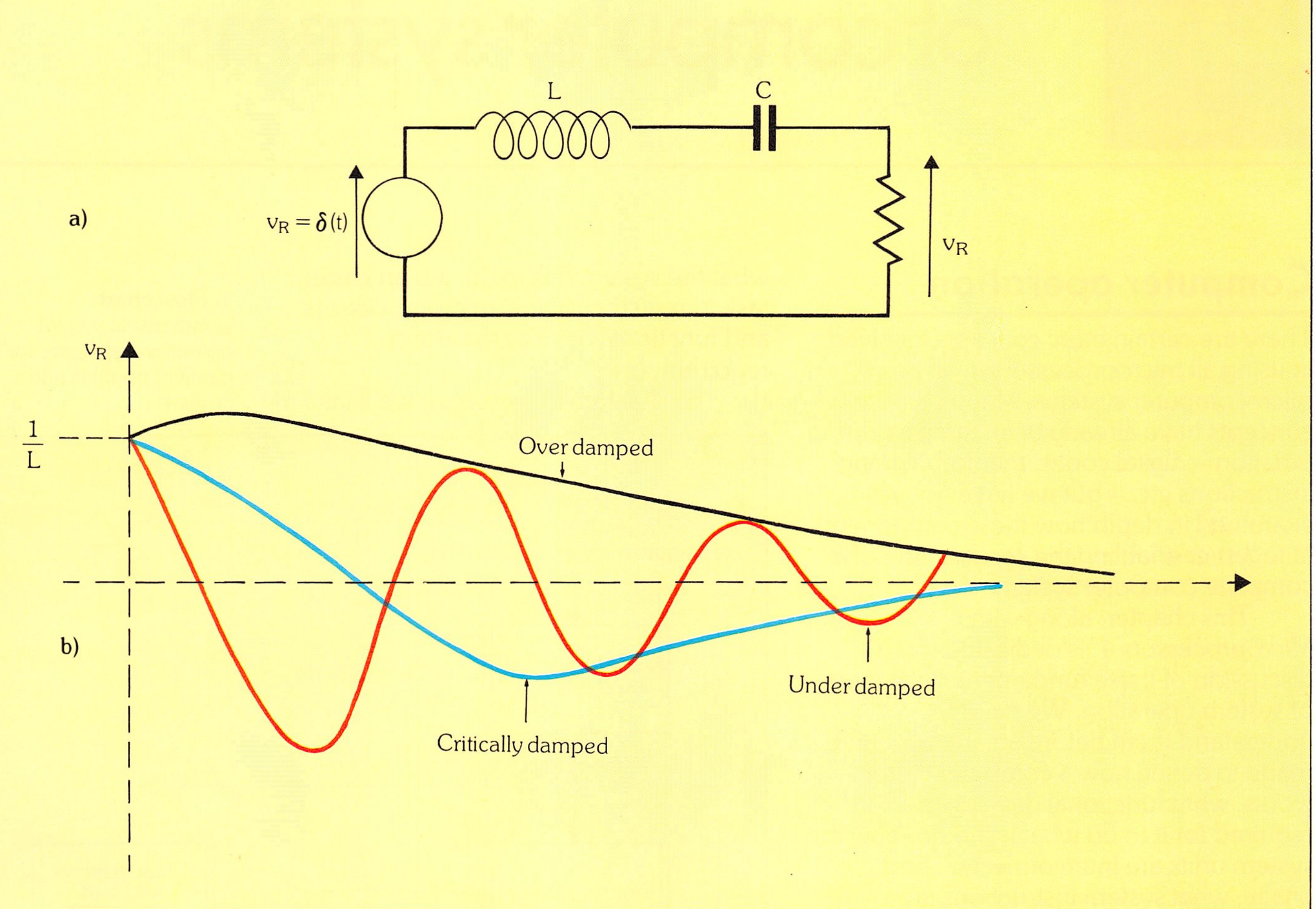
Two port networks

Looking now at one or two simple two-port networks, *figure 3a* shows a resistor and capacitor connected in series with an impulse voltage generator, and the output voltage is taken across the capacitor.

We may find the output voltage, v_o, by first approximating the input voltage by a pulse of finite duration as in *figure 3b*. In an earlier *Basic Refresher* we studied this circuit when it was excited by a step of voltage and the pulse which we are now using may be considered as a positive step of size $1/t_1$ at time $t = 0$, followed by an equal negative step at time $t = t_1$. The voltage across the capacitor in response to each of these separately is shown by the



4



2. A unit impulse voltage generator connected to: **(a)** a resistor; **(b)** an inductor; **(c)** a capacitor.

3. A unit impulse voltage connected to a two-port network.

4. A series RLC circuit excited by an impulse of voltage.

black curves of *figure 3c*; and the total response is shown by the red curve, by adding together the responses to the two separate steps.

As we further reduce the length of the pulse, keeping its area constant, we ultimately obtain the response to a unit impulse as shown in *figure 3d*. Here we see that the capacitor voltage rises suddenly to a value of 1/RC, and then falls in an exponential manner with a time constant of 1/RC.

In a similar way we could show that the voltage v_R taken across the resistor will consist of an impulse at $t = 0$, followed by a negative exponential curve as shown in *figure 3e*.

Response of a tuned circuit

As a final example, let us now consider the current flowing in a series RLC circuit excited by an impulse of voltage. First, we approximate the impulse by a finite length pulse. If we look at the voltage across the resistor of *figure 4a*, we get the response voltage v_R as shown in *figure 4b*. The form of the output is determined by the value of the resistance R: if it is small, the circuit is underdamped and the response is shown by the red curve; if it is large, the circuit is overdamped and the response is as the black curve, settling towards zero without any change of sign. For the critically damped case where:

$$R = 2\sqrt{\frac{L}{C}}$$

the response is as shown by the blue curve, and there is a single change of sign before settling to the final value, and no continuous oscillation. □

MICROPROCESSORS

Fundamentals of computer systems

Computer operation

There are certain basic concepts involved in using all microprocessors in all microcomputer systems. Many of these concepts have already been discussed in isolation – digital codes, memory, timing, instructions etc. – but we have not yet examined in depth how these components fit together enabling the construction of a complete computer system.

This chapter, along with *Microprocessors 4*, is dedicated to a discussion of these fundamental concepts of system operation. We need to understand them before an attempt can be made to define how a computer system works; what functional devices are required for it to do what it should; how the system units are interconnected; and finally, what system instructions are required.

System evaluation

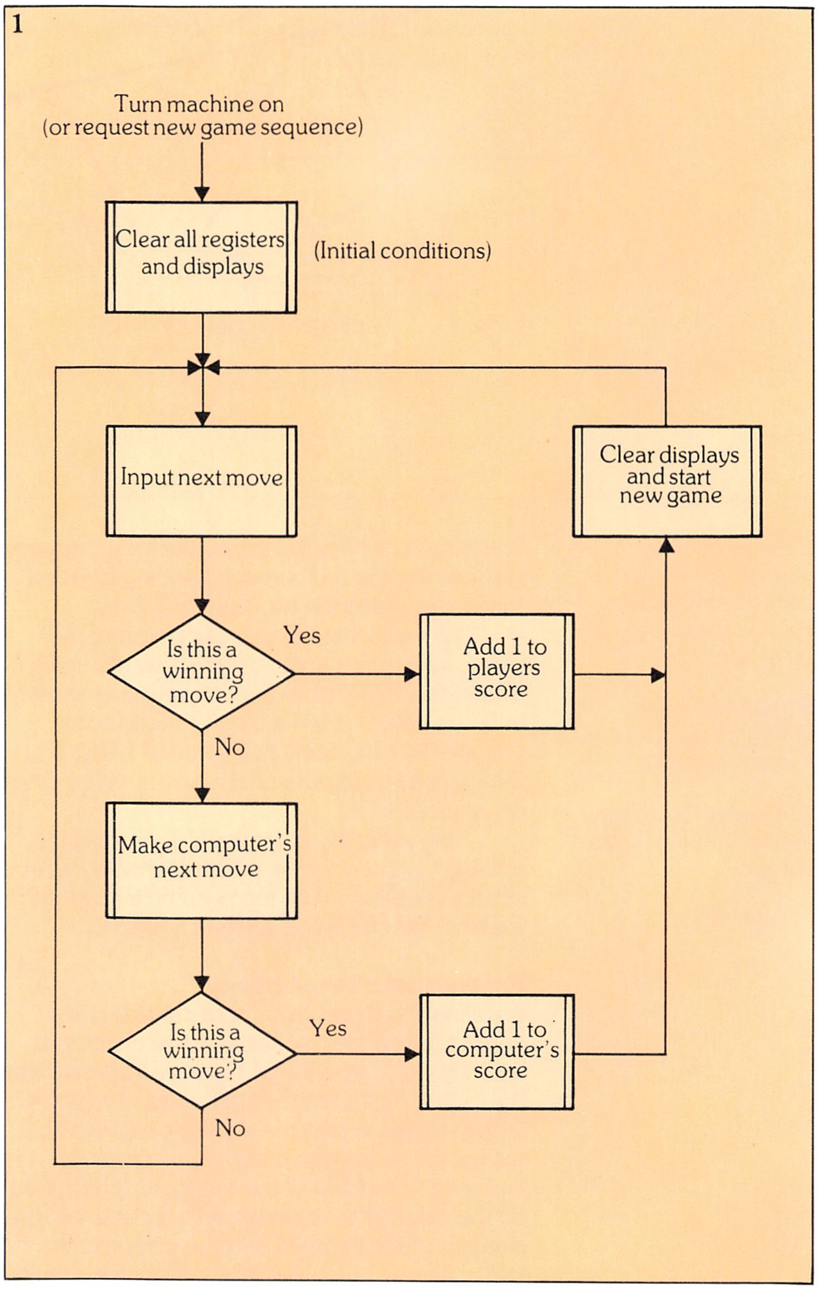
The first step to be taken is to define, in broad outline, a system definition. This process may be called **system evaluation**. System evaluation often begins with a definition of the features available in the microprocessor or microcomputer. Examining these features, along with the requirements of the complete system, helps to determine what functional devices will perform all the required tasks. The program, i.e. the step-by-step sequence of controlling instructions, is not, incidentally, considered at this stage.

Selecting system units

System evaluation generally begins with a complete definition of what the system must do in its final form. Often aids, such as lists or flow charts showing the various steps and functions of the system, are used to help define the system and its order of events. With a complete understanding of what the system has to do, it is an easier task to match available microprocessors and functional devices to system requirements.

To illustrate this process, we'll take a

1. Flow chart summarising system operation necessary for a game of noughts and crosses.



simple familiar example and develop such a description. Consider a microprocessor-based system that plays the game noughts and crosses with a human partner. Noughts and crosses, remember, needs two players who each, in turn, draw either an X or an 0 in any empty square; the 'board' comprises 9 squares in a 3x3 matrix. The player who chooses X must thereafter only draw Xs throughout the game; similarly for the player who chooses 0s. The first player to draw three Xs or 0s in a line, horizontally, vertically, or diagonally, is the winner of the game. Obviously, to enable a microcomputer system to play, some sort of game board must be used, which shows the player's moves and the state of the game after each move.

Table 1
Microprocessor features related to different applications

	Applications requirements		
Microprocessor features	**Simple controllers**	**High-speed communications**	**Accurate arithmetic**
Bit length	4	8 or 16	16 or 32
Speed	Low	High	High
Instruction set needs	Simple	Efficient data transfers and comparisons	Efficient arithmetic operations
Timing/control	Simple	Simple	Simple
Interrupt structure	Simple, slow	High speed multiple level	Multiple level

Initially, the system is in a condition where it is waiting for someone to begin the game; it must then recognise that the game has begun, by accepting an X or 0 as input information, as someone takes a turn. The microprocessor must then check to see if that move has won the game. The system must then make a move, and check to see if its move has won the game. The game continues until a move wins the game. As the series of games progresses, the system must keep track of the score, incrementing each score by one, as appropriate. Finally, after each game has been won, the system must clear the board in preparation for the next game.

Operation list or flow chart

We have seen how the game operates, and we have been able to describe it in quite straightforward terms. However, it may be summarised more clearly as a list of operations:

1) The system is turned on. This causes it to commence at the first game of a series of games, waiting for someone to make a move. A method must also be provided to cause the system to start a new series of games if games have already been played.
2) The person's move is received and displayed on the game board.
3) The system checks to see if the person's move has won the game. If it has, one is added to the player's score, and a new game is begun.
4) If the game is not won, the system makes a move.
5) If the system's move wins the game, the system adds one to its score and starts a new game.
6) If there is no winner at this stage, the system goes back to step 2 and continues.
7) A new game is started by clearing the game board, displaying the scores, and going back to step 2 to continue play.

The alternative method of summarising system operation is a flow chart as in *figure 1*. Rectangular blocks are used in the flow chart to represent processes which the system must perform, and diamond shaped blocks are used to represent decisions which the system must make. A flow chart provides a graphic overview of what the system must do, and how the various tasks it performs relate to one another.

Whether a flow chart, list, or both methods are used, depends on the personal choice of the system designer.

Choosing a microprocessor

The main features which are important in the choice of a microprocessor for a specific application are: bit length; speed; its instruction set; the timing and control signals available which are used to control other functional devices; its interrupt procedure.

All of these features will, in some way, be related to the task to be performed: some may be important and some may not. *Table 1* considers these features in

relation to some specific applications.

Bit length

The **bit length** of the microprocessor refers to the number of data bits which may be handled at one time. As we know, microprocessors can handle data in 4, 8, 16 or 32-bit words.

For simple tasks, such as our noughts and crosses game, where only limited accuracy is needed and only a small number of different instruction codes are required, a 4-bit microprocessor is more than adequate. Tasks which need a greater level of accuracy, or where long strings of number character codes are involved, require a 16 or even 32-bit microprocessor.

The number of different system conditions that any microprocessor can handle directly can be calculated very simply: if, for example, the microprocessor is an N-bit device, then it can handle 2^N different system conditions. A 4-bit microprocessor can therefore handle up to 2^4, or 16 different conditions; an 8-bit microprocessor can handle 256 (2^8) conditions; a 16-bit microprocessor 65,536; and finally, a 32-bit microprocessor can handle the staggering sum of 4.29×10^9 conditions, directly.

The numerical accuracy which an N-bit microprocessor is capable of attaining in a single operation is $1/2^N$ or, as a percentage, $100/2^N$. Thus, a 4-bit microprocessor is capable of directly attaining an accuracy of:

$$\frac{100}{16} \approx 6\%$$

i.e. the result is within 6% of the actual, correct calculation. Accuracies of 0.4%, $1.5 \times 10^{-3}\%$ and $2.3 \times 10^{-8}\%$ are attainable with 8, 16 and 32-bit microprocessors respectively.

Any microprocessor may be used indirectly to any required accuracy by simply processing more bits, a word at a time. So, for example, a 4-bit microprocessor may give the same accuracy as a 16-bit microprocessor by processing 4, 4-bit words for every one, 16-bit word which the 16-bit microprocessor processes. Obviously, however, this is done at the expense of overall operating speed.

Speed

Besides bit length, the speed at which a microprocessor is capable of handling a given task is related to other factors, such as the power of the instruction set, timing and control procedures, interrupt procedures, and the response of related functional devices (memory input/output devices etc).

Speed is important when sending and receiving information. It is also important when performing difficult and complex computations. In both of these cases, a microprocessor which offers high bit-length operation with fast clock speeds and an efficient instruction set is required. For our noughts and crosses system, on the other hand, where the operations are simple and the second player is a person (who is a slow opponent anyway – at least, compared to even the simplest microprocessor), speed is almost totally unimportant.

Instruction set

As we know, a microprocessor follows steps in a program. These steps (and therefore the program itself) are defined in terms of the instructions which a microprocessor recognises. The instructions which one type of microprocessor recognises and follows may not be the same instructions as recognised and followed by another. So, a close check must be made before choosing a microprocessor to ensure that the instructions in its instruction set are sufficient to handle the task required.

A calculator-type of microprocessor system, for example, needs efficient arithmetic operations including addition, subtraction, multiplication, division, exponents, and so on. A communications system, on the other hand, needs a wide choice of input and output instructions that can transfer information rapidly.

A logical system, say, a microprocessor-based chess game, needs extensive logical, comparative and decision-making instructions. A simple system needs only a very small set of instructions.

(continued in part 40)